COLOUR APPEARANCE MODELLING OF UNRELATED SELF-LUMINOUS STIMULI

Martijn WITHOUCK

Members of the Examination Committee:
Prof. B. Meesschaert, chair
Prof. P. Hanselaer, promotor
Dr. K. Smet, copromotor
Prof. G. Deconinck, copromotor
Prof. W. Ryckaert, assessor
Prof. A. Heylighen, assessor
Prof. J. Wagemans, assessor
Dr. E. de Beer, assessor
Prof. Y. de Kort
Dr. M. Tarpan

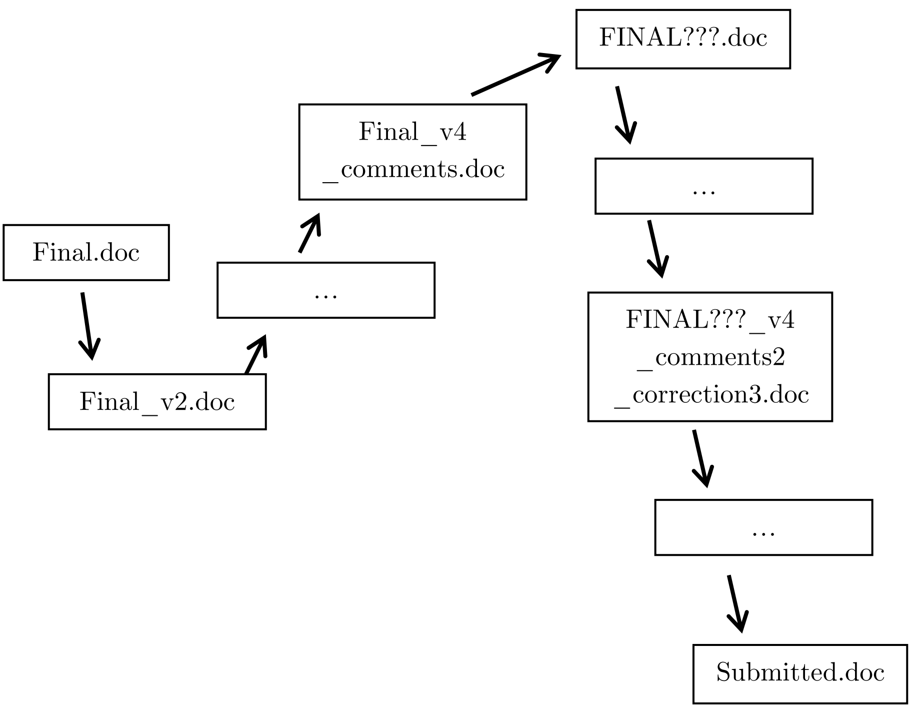
Dissertation presented in partial fulfilment of the requirements for the degree of PhD in Engineering Technology.

June 2015

Uitgegeven in eigen beheer, Martijn Withouck, Gebroeders De Smetstraat 1, B-9000 Ghent (Belgium)

ACKNOWLEDGEMENTS

Completing a PhD is more than explaining (or trying to explain) to relatives what the job is all about. It is a long journey with several mountains to be climbed. Beginning with a long literature study (let's call that one the Mont Blanc) and ending with the writing of this thesis (the famous K2 mountain). Along the way, several other mountains, some with dangerous valleys in between, were encountered. Publishing the results of an investigation as a journal paper can be represented by the struggle getting out of such a valley through the main river, where one thinks to have finished the paper written down in a *Final.doc* document (represented by reaching the main river) but in reality finds himself in a new series of corrections, for example: *Final_ v4_ comments.doc* (represented by one of the many river deltas). The struggle of finding the paper draft where the supervisors gave up giving comments (*Submitted.doc*), is illustrated in the figure below.



Along this challenging journey, lots of people kindly helped. First of all I want to thank Peter Hanselaer and Wouter Ryckaert for taking me on board the enthusiastic, driven Light&Lighting Laboratory. Together with Kevin

Smet, they supported me in every step of the PhD. Their attention to detail and hard work were motivating, convincing me to continue keeping up the good work, improving the predictive power of our models,... Many thanks!

Furthermore, I am very grateful to Geert Deconinck, Esther de Beer, Ann Heylighen and Johan Wagemans for their support. Their insightful comments and suggestions, based on an extensive expertise of several research topics, broadened my view on colour appearance research, leading to key adjustments of the scientific approach taken.

Furthermore, I deeply appreciate the time and effort the members of my examination committee have put into the evaluation of my dissertation.

I would also like to thank the Research Council of KU Leuven for supporting this research project (STIM-OT/11/056 and OT/13/069).

My gratitude is also extended to Mike Pointer. Being able to consult Mike's giant knowledge of colour appearance was critical to my research. Time after time, I received an answer to my questions within a couple of hours. In addition, I have been privileged to get to know Steve Fotios. Steve gave great constructive comments after my first conference presentation and gave me the opportunity to contribute to a CIE publication about brightness.

I am also very grateful to all observers taking part in one of the many psychophysical experiments. Seduced with candy and soft drinks, they were very dedicated (but at times also very bored) in participating my experiments.

In addition, I have had the joy to work with many colleagues and students in last five years. Thank you all for your contribution to my research and the atmosphere at the lab, and for laughing with my overdose of stupid jokes.

Of course no acknowledgments would be complete without giving thanks to my parents. They have had a big role in who I am and taught me about hard work and how to be independent.

Last, but certainly not least, I must acknowledge, with tremendous and deep thanks, my wife, Caro Cools. Caro, thank you for your support, love, encouragement, listening ear,...

Gent, June 2015 Withouck Martijn

ABSTRACT

Colour appearance models, i.e. models that attempt to predict the colour appearance of a stimulus by taking the physical properties of the stimulus and its surroundings into account, have been developed and investigated for more than 40 years. Most of these models were developed to handle related colours, i.e. colours perceived in relation to other colours. A typical example is the 'reflected' colour of an object as seen in an illuminated scene. However, two models - CAM97u and CAMFu - were developed to predict the appearance of unrelated colours, i.e. colours perceived in isolation from any other colour (e.g. a traffic light seen at night). Unfortunately, due to the lack of psychophysical data, neither of these two models has been investigated extensively. Before being able to extend these models to other types of stimuli and viewing conditions, they need to be validated using new visual data.

The aim of this doctoral research is *to investigate the colour appearance of unrelated self-luminous stimuli*. An accurate prediction of the colour appearance of these stimuli through a colour appearance model can be a valuable tool: it can assist in the development of requirements for light-emitting diode (LED) signs, in the standardization of the appearance of marine, aviation or traffic lights viewed during a dark night, in the continuous development of colour appearance models for other viewing conditions,…

In a first series of psychophysical experiments, the brightness of stimuli with a constant luminance has been evaluated by a group of observers. The stimuli were shown in a darkened room, specially designed for this doctoral research project. In the centre of one wall, a circular self-luminous area was present. The colour of this stimulus area was computer controllable by adjusting the flux of the R(ed)G(reen)B(lue)W(white) LED behind it. The observers viewed the stimulus area from a distance that ensures a 10° field of view. The brightness evaluation of these stimuli was performed using a magnitude estimation method by scaling the brightness of each test stimulus compared to that of a reference stimulus to which a brightness value of 50 was attributed. The predictive performance of the CAM97u and CAMFu colour appearance models and four other vision models, specially designed to predict *brightness*, was investigated. Due to, among others, a severe underestimation of the effect of colourfulness on brightness - also known as the Helmholtz-Kohlrausch effect - none of the models seemed to be able to adequately predict the brightness perceived by the observers. Adapting the CAM97u model by increasing the colourfulness contribution in the brightness attribute, resulted in a modified model, called CAM97um, which allows for a substantially better brightness prediction. The performance of this new model was confirmed by the results of both a matching experiment and an extensive magnitude estimation experiment in which the test stimuli covered a wide range of luminance and chromaticity values.

In a subsequent series of psychophysical experiments, in addition to the brightness, the hue and "amount of white" perception of unrelated self-luminous stimuli was also investigated using a magnitude estimation method. The amount of white is a newly proposed attribute, and basically corresponds to a layperson's conception of attributes such as colourfulness, chroma or saturation. It was introduced based on the results of a preliminary pilot study revealing that laypersons often have difficulty understanding, and hence judging, the colourfulness of a stimulus. Again, unrelated self-luminous 10° stimuli, with a wide range of luminance and chromaticity values, were evaluated by observers in the darkened room. Based on the obtained visual data, a new colour appearance model for unrelated self-luminous stimuli, CAM15u, was developed. The main features of the model are the use of the absolute spectral radiance of the stimulus as input, the use of the CIE 2006 cone fundamentals and a simplified calculation procedure compared to existing models. The model predicts the brightness, hue, colourfulness, saturation and the amount of white. The CAM15u model is restricted to photopic, non-glare-inducing unrelated stimuli having a field of view of 10°. The model was validated using the results of an additional experiment. It was found that, despite its simplicity, CAM15u performs as well or better than other, more complicated, CAMs.

In a final series of psychophysical experiments, the brightness perception of different sized, unrelated self-luminous stimuli was investigated in a magnitude estimation experiment. The stimuli were shown in a darkened room on a wide gamut LCD monitor. A significant, hue independent, effect of stimulus size on brightness was found, effectively modeled by a simple power function. Finally, the dependence of brightness on stimulus size was incorporated into the brightness prediction of the CAM15u model. The predictive performance of the modified brightness prediction was validated using the results obtained in an additional experiment in which observers evaluated the brightness of unrelated self-luminous test stimuli with variable size, chromaticity and luminance.

Although further improvements and extensions are still possible, CAM15u has proven its value in predicting the appearance of unrelated self-luminous stimuli. It can be a valuable tool for the improvement of existing standards and guidelines for traffic signs, LED billboards,...

SAMENVATTING

De hoofddoelstelling van dit doctoraatsproject is het modelleren van het kleuruitzicht van ongerelateerde lichtgevende kleuren. Ongerelateerde kleuren zijn kleuren waargenomen in isolatie van andere kleuren, zoals een verkeerslicht geobserveerd in nachtelijke omstandigheden. Een accurate voorspelling van de perceptie van dergelijke stimuli a.d.h.v. een 'colour appearance' model kan een waardevolle bijdrage leveren in de ontwikkeling van richtlijnen voor led signalisatie en reclameborden, in de standaardisatie van verkeerslichten en borden met variabele boodschap, in het beschrijven van de verblinding van verlichtingstoestellen en voor de ontwikkeling van nieuwe specifieke 'colour appearance' modellen. Tijdens de voorbije 40 jaar werden verschillende 'colour appearance' modellen ontwikkeld en onderzocht. Deze modellen trachten het kleuruitzicht of de 'colour appearance' van een stimulus te voorspellen a.d.h.v. de fysische eigenschappen van de stimulus en zijn omgeving. Het merendeel van deze modellen werd ontwikkeld voor 'gerelateerde kleuren', dit zijn kleuren waargenomen in relatie met andere kleuren. Een typisch voorbeeld hiervan is de gereflecteerde kleur van een object gezien in een verlichte omgeving. Daarnaast werden o.a. twee modellen ontwikkeld die de perceptie van 'ongerelateerde kleuren' trachten te voorspellen, CAM97u en CAMFu. Door een gebrek aan psychofysische data werden deze bestaande modellen echter nog nooit grondig gevalideerd. Vooraleer deze modellen uit te breiden naar complexere stimuli en andere kijkomstandigheden, leek het dus aangewezen om ze eerst te verifiëren m.b.v. nieuwe visuele data.

In een eerste reeks psychofysische experimenten werd de helderheid van een reeks lichtgevende teststimuli omgeven door een zwarte achtergrond, met een constante luminantie maar variabele kleurtint en saturatie, geëvalueerd door een groep waarnemers. Deze stimuli werden gegenereerd d.m.v. rode, groene, blauwe en witte leds gepositioneerd achter een cirkelvormige diffusor. Door de flux van deze leds te variëren kan de kleur van dit lichtgevend oppervlak aangepast worden. De afstand van de waarnemer tot de stimulus werd zodanig gekozen dat het gezichtsveld van de stimuli ongeveer 10° bedroeg. Via de 'magnitude estimation' methode werd de helderheid van de teststimuli geëvalueerd in vergelijking met die van een referentiestimulus waaraan een helderheidswaarde van 50 werd toegekend. De mate waarin CAM97u, CAMFu en vier andere kleurmodellen de waargenomen helderheid

kunnen voorspellen werd onderzocht. De overeenkomst met de visuele data was echter voor alle modellen pover tot matig, onder andere door een onderschatting van het effect van de kleurrijkheid of 'colourfulness' van een stimulus op zijn helderheid. Dit effect staat bekend als het Helmholtz-Kohlrausch effect en heeft als gevolg dat gesatureerde, levendige stimuli er steeds helderder uitzien dan de meer neutrale kleuren (voor dezelfde luminantie). Door in CAM97u de bijdrage van de kleurrijkheid te verhogen, geformaliseerd in een aangepast model CAM97um, werd een substantieel betere voorspelling van de helderheid bekomen. De helderheidsvoorspelling van dit CAM97um model werd zowel via een 'matching' als via een uitgebreid 'magnitude estimation' experiment gevalideerd.

In een daaropvolgende reeks van psychofysische experimenten werd, naast helderheid, ook de perceptie van kleurtint en 'hoeveelheid wit' onderzocht. De 'hoeveelheid wit' is een nieuw kleurattribuut dat werd geïntroduceerd na het uitvoeren van enkele preliminaire experimenten. Hiermee werd geprobeerd om de beoordeling van de traditionele kleurattributen kleurrijkheid, saturatie en chroma door personen met weinig of geen kennis over kleurtheorie te vereenvoudigen. Opnieuw werden in de psychofysische experimenten ongerelateerde lichtgevende 10°-stimuli, dit keer met een variabele luminantie, kleurtint en saturatie, getoond aan proefpersonen. Gebaseerd op de resultaten van deze visuele experimenten werd een nieuw 'colour appearance' model voor ongerelateerde lichtgevende stimuli ontwikkeld, CAM15u genaamd. Het model gebruikt de absolute spectrale radiantie van de stimulus als input, maakt gebruik van de CIE 2006 'cone fundamentals' en voorziet in een vereenvoudigde berekeningsmethode vergeleken met de bestaande modellen. Het tracht de helderheid, kleurtint, kleurrijkheid, saturatie en hoeveelheid wit te voorspellen. Het CAM15u model is beperkt tot fotopische, niet-verblindende ongerelateerde 10°-stimuli. Via een uitgebreid validatie-experiment werd aangetoond dat, ondanks zijn eenvoud, de voorspellingen van CAM15u gelijkwaardig of zelfs beter zijn dan deze van de andere modellen.

In een laatste reeks van psychofysische experimenten werd de helderheidsperceptie van ongerelateerde lichtgevende stimuli met variabele grootte, luminantie, kleurtint en saturatie onderzocht. De stimuli werden gepresenteerd op een lcd-scherm in een donkere ruimte. Een significant, kleurtint-onafhankelijk, effect van de stimulusgrootte op helderheid werd waargenomen. De impact van de stimulusgrootte werd gemodelleerd d.m.v. een eenvoudige machtsfunctie en dit werd opgenomen in de helderheidsvoorspelling van CAM15u, resulterend in de uitgebreidere versie

CAM15us (waarbij de toegevoegde “s” staat voor “size”). De resultaten van een validatie-experiment toonden aan dat CAM15us in staat is om de helderheidsperceptie van ongerelateerde lichtgevende stimuli, met variabele grootte, chromaticiteit en luminantie, te voorspellen.

Hoewel er nog steeds ruimte is voor verbetering en uitbreiding, hebben CAM15u en CAM15us zeker hun waarde bewezen in het voorspellen van de kleurperceptie van ongerelateerde lichtgevende stimuli. Daarmee is het pad geëffend om ook gerelateerde stimuli, zoals stimuli omgeven door een lichtgevende achtergrond, te modelleren.

LIST OF ABBREVIATIONS

ANOVA	analysis of variance
ANCOVA	analysis of covariance
ASTM	American Society for Testing and Materials
CAM	colour appearance model
CI	confidence interval
CIE	Commission International d'Eclairage
CMF	colour matching function
CV	coefficient of variation
DMX	digital multiplexed
EEW	equi-energy white
FOV	field of view
H-K effect	Helmholtz-Kohlrausch effect
JND	just noticeable difference
LCD	liquid-crystal display
LED	light-emitting diode
NCS	Natural Colour System
NSVV	Nederlandse Stichting Voor Verlichtingskunde
ROA	Richtlijnen voor het Ontwerpen van Autosnelwegen
SPSS	Statistical Package for the Social Sciences
VAC	variable-achromatic-colour method
VCC	variable-chromatic-colour method

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

In my second year as doctoral researcher, I attended my first conference meeting abroad. After the first day - and some beers - I took a taxi towards my hotel. Surprisingly, at the first traffic lights, the taxi driver disregarded the red light. Terrified, I asked for an explanation. "Don't worry Sir, I have a lot of experience. I've learned driving from my brother, he never stops at red lights." Shocked by the situation and on the edge of my seat, I stayed in the car. Of course, a few minutes later, we approached the next traffic lights. Before being able to jump out of the car, the lights turned green and curiously the car stopped. The driver noticed my confused gaze and started laughing: "Sir, don't worry. I'm not crazy! My brother often drives up here…"

Investigating the colour appearance of unrelated self-luminous stimuli, like a traffic light seen at night, does not include an evaluation of a driver's response when seeing the traffic lights turning red. What this doctoral research project is about, and what is meant by an unrelated self-luminous stimulus and much more is described in this chapter.

1.1 Introduction

The word 'colour' has different meanings according to its use: the colour of an object (hue, saturation,…), the 'true colours' of somebody (as in the song of Cindy Lauper), the 'colours of the wind' in Pocahontas (Walt Disney film), the expression 'bringing colour into your life',… When 'colour conference' is searched on the world wide web, the first result is the 'Colour Conference' for women of all ages, backgrounds and cultures trying to make the world a better place. Of course, colour mainly has a perceptual meaning. The Oxford Dictionaries describe colour as the property possessed by an object producing different sensations on the eye as a result of the way it reflects or emits light. The CIE (Commission International d'Eclairage) defines colour as the characteristic of visual perception that can be described by attributes of hue, brightness (or lightness) and colourfulness (or saturation or chroma) [1].

Colour has intrigued a lot of people throughout the past centuries. Many scientists - Isaac Newton, Thomas Young, Ewald Hering,… - proposed theories about colour, while artists like Claude Monet, Vincent Van Gogh, Pablo Picasso,… used these colour theories to create brilliant paintings. Today, these theories of colour are still important, e.g. to reproduce colours for photography, medical imaging, coloured displays,…

Reproducing colours is however a severe challenge as several colour phenomena occur when a complex scene is being viewed. The appearance of the scene is influenced by the surround condition (bright, dim, coloured,…), the contrast between colours in the scene, the state of adaptation of the human visual system, the texture of the object,… Research into and the modelling of the colour appearance of such complex scenes is a big challenge and has a long history. However, even after several decades no such model, called a 'colour appearance model', yet exists that is capable of treating complex scenes. Only colour appearance models (abbreviated as CAMs hereafter) developed for simple viewing conditions are available.

In this doctoral research project, the colour appearance of unrelated self-luminous stimuli has been investigated. Unrelated colours are colours perceived to belong to an area seen in isolation from other colours [1]. A typical example of an unrelated colour is a self-luminous stimulus surrounded by a dark background, like a marine or traffic signal light viewed during a dark night. Due to the absence of other colours and a real luminous background, the description of the perception of these stimuli can be considered as being relatively simple and elementary compared to stimuli seen in relation to other stimuli. During the following chapters, it will become clear that ‘relatively simple’ is not quite simple at all. Experimental setups and methods have been developed to investigate the colour appearance of unrelated self-luminous stimuli. Based on the results of these visual experiments, several CAMs developed to handle unrelated colours have been investigated and a new colour appearance model, CAM15u, has been designed.

This chapter shortly introduces the basic mechanisms of colour vision and the concept of colour appearance modelling.

1.2 Colour vision

Human colour vision starts with light absorption by the photo-sensitive receptor cells in the retina. Light is the part of radiation able to excite the human visual system, with a wavelength roughly between 380 nm and 780 nm. Two kinds of receptor cells can be distinguished, the rods and the cones. The rods, mainly responsible for scotopic vision, are sensitive to low intensity visible radiation (luminance below 5 cd/m^2). The cones, dominating photopic vision (luminance above 5 cd/m^2), come in three different types and are typically referred to as the ρ, γ, β cones, with peak sensitivities located around 569 nm, 541 nm and 448 nm, respectively (see Fig.1.1). These cones are also denoted with other symbols such as *LMS* or *RGB*, suggestive of long-, middle-, and short-wavelength or red, green, and blue sensitivity.

The specific shape of the spectrum - the amount of power present in the physical stimulus at each wavelength - has an influence on the perceived colour due to the different sensitivity of these three cone types. The compressed responses of these cones, resulting from the light absorption, are transformed into an achromatic and two opponent signals (red-green and yellow-blue) and sent to the brain through the optic nerve (see Fig.1.1). Finally, the colour perception is the result of further processing of these signals in the brain. This further (cortical) processing is dependent on the spectrum of the stimulus and surround, the state of adaptation of the human visual system, our memories of previous perceived stimuli,… Furthermore, when investigating the colour appearance of stimuli also semantics come into play.

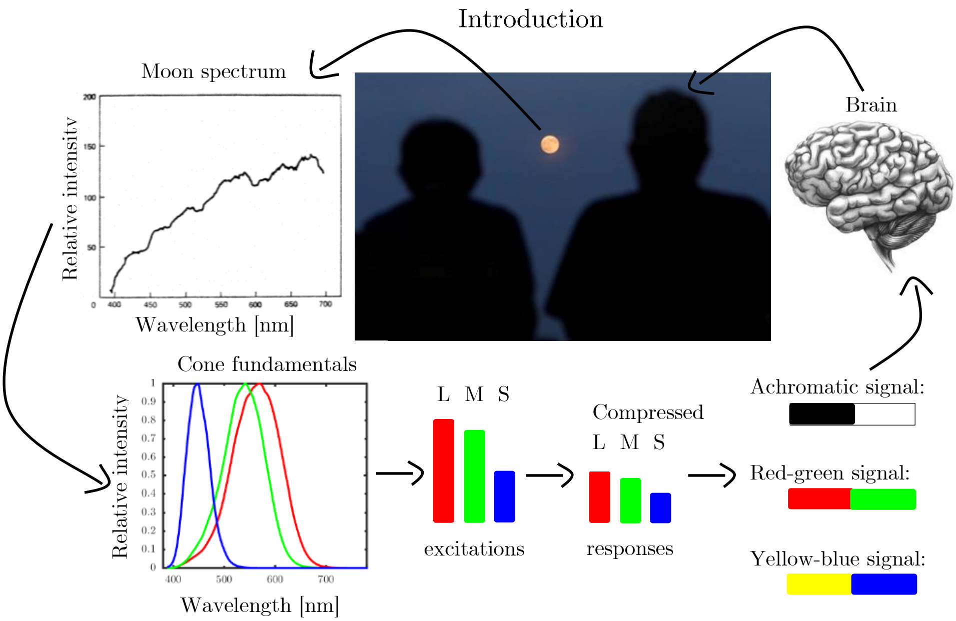


Fig.1.1. A brief overview of the human colour vision.

1.3 Colour appearance modelling

Colour appearance is defined by the CIE as the aspect of visual perception by which things are recognized by their colour [1]. This visual perception depends on the spectral aspects of both the visual stimulus and the viewing conditions.

1.3.1 Types of stimuli and viewing conditions

Depending on the type of stimulus and the viewing conditions, the perceived colour may appear in several modes of colour appearance. Below, the CIE definitions of different colour appearance modes are given [1]:

Aperture colour: perceived colour seen through an aperture - an opening that defines the area over which average optical emission is measured - which prevents its association with a specific object or source (for example: a colour perceived as filling a hole in a screen – for which there is no definite spatial localisation in depth).

Corresponding colours: pairs of colour stimuli that have the same colour appearance when one is seen in one set of adaptation conditions and the other is seen in a different set.

(Self-)Luminous colour: colour perceived to belong to an area that appears to be emitting light as a primary light source, or that appears to be specularly reflecting such light (with an angle of reflection equal and opposite to the angle of incidence).

Non-luminous colour: colour perceived to belong to an area that appears to be transmitting or diffusely reflecting light as a secondary light source.

Object colour: colour perceived as belonging to an object.

Related colour: colour perceived to belong to an area seen in relation to other colours.

Surface colour: colour perceived as belonging to a surface from which the light appears to be diffusely reflected or radiated (the light is deviated in many directions).

Unrelated colour: colour perceived to belong to an area seen in isolation from other colours.

Volume colour: colour perceived as belonging to the bulk of the substance

1.3.2 Perceptual attributes

The colour perception can be described by absolute and relative fundamental, univariate perceptual attributes. The absolute attributes are:

Brightness: the attribute of a visual perception according to which an area appears to emit, or reflect, more or less light.

Colourfulness: the attribute of a visual perception according to which the perceived colour of an area appears to be more or less chromatic.

Hue: attribute of a visual perception according to which an area appears to be similar to one of the following colours: red, yellow, green, and blue, or to a combination of adjacent pairs of these colours considered in a closed ring.

Based on the absolute attributes brightness and colourfulness, the relative attributes are obtained:

Chroma: colourfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears white or highly transmitting (only for related colours).

Lightness: brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting (only for related colours).

Saturation: colourfulness of an area judged in proportion to its brightness.

1.3.3 Colour appearance models

The goal of any CAM is to provide a full description of the colour appearance of a stimulus in terms of the perceptual attributes as mentioned above under a variety of viewing conditions. In addition to the physical measurements of the stimulus, measurements of the prevailing viewing conditions are used as input to the models. Its development typically follows a stepwise approach: a basic model is extended step-by-step to include more complex conditions and various colour appearance phenomena.

Over the last forty years, various colour phenomena have been investigated and several CAMs have been developed, one more complex than the other. As a result of the work by, amongst others, Hunt [2-4], Nayatani [5, 6], Fairchild [7, 8] and Luo [9, 10], the CIE proposed a simplified colour appearance model, CIECAM97s [11] for *related colours*. Five years later a new model, based on CIECAM97s, was developed: CIECAM02 [12]. Although widely used it is still subjected to improvements [13, 14].

While much colour appearance research has focussed on related stimuli, the research on unrelated colours is rather limited. Based on a number of different versions of his CAM for related colours, Hunt developed CAM97u, a model for *unrelated colours* [15, 16]. More recently, Fu et al. has presented another model for unrelated colours, called CAMFu, based on CAM97u and CIECAM02 [17].

1.3.4 Applications of CAMs

As mentioned earlier, colour appearance research and modelling has a long history. Driven mainly by its importance for practical application, a great deal of progress was made from the late sixties to the eighties, by the Kodak Research Laboratories. Back then, photography was - together with colour television - dominating the colour image reproduction industry. In the quest for improved image quality of prints, among others Robert Hunt and Michael Pointer developed several models eventually leading to the CIECAM97 and CIECAM02 models. Today imaging science is more than photography and colour television; other applications such as digital cameras, digital graphics, medical imaging, forensic imaging, hardcopy, colour display devices,… come into play and they continue to provide a driving force for the development of new colour appearance models. As mentioned above, developed by the CIE and still subjected to improvements, CIECAM02 is undoubtedly the most famous and widely used CAM. The model is used in a plugin for Adobe Photoshop [18], in Android applications as a colorimeter [19], in the colour management software of Hewlett Packard printers [20], CANON devices and Microsoft Windows operating systems [21],…

1.4 Colour appearance of unrelated self-luminous stimuli

1.4.1 Goal of the doctoral research project

As mentioned earlier, a CAM for unrelated colours was designed more than 15 years ago: CAM97u [15, 16]. At first, the goal of this doctoral research was to develop a new CAM for self-luminous stimuli viewed on both a dark (cfr. unrelated colours) and a self-luminous (cfr. related colours) background. Starting with visual experiments using a set of *unrelated self-luminous* stimuli, CAM97u would be validated. Later on, CAM97u would be extended to work with self-luminous stimuli viewed on a luminous background by using the principles of CIECAM02 for related colours and the results of visual experiments with *self-luminous stimuli surrounded by a luminous background*. Such a model would fill an existing gap in colour appearance modelling as none of the existing CAMs is able to handle self-luminous stimuli. Typically, these stimuli are seen in relation with the background but are, in contrast to related colours, not similarly influenced by the illumination. For example, a green traffic light can be perceived in relation with a building illuminated by the sun or by the streetlights, but the spectrum of the traffic light is not correlated with the spectrum reflected from the building. Nonetheless, the perception of the stimulus would still be influenced by the different physical properties of the background. In fact, a new appearance mode could be defined:

Correlated colour: colour perceived to belong to an area seen in relation to another colour, whereby the physical colour quantities are correlated, e.g. due to a common illumination.

Uncorrelated colour: colour perceived to belong to an area seen in relation to another colour, whereby the physical colour quantities are uncorrelated: e.g. the colours are illuminated independently or, one or the other (or both) is self-luminous.

Summarised, the goal of the doctoral research was to develop a CAM for *uncorrelated self-luminous* colours. However, the CAM97u model was never verified experimentally [22] and preliminary benchmark experiments soon indicated CAM97u was unable to accurately predict the perceived brightness of unrelated stimuli [23]. Gradually, as a result of this poor predictive power of CAM97u, the goal of the doctoral research project was changed to 'developing a CAM for *unrelated self-luminous* stimuli'. As mentioned earlier, a colour appearance model attempts to predict the colour appearance - the visual perception by which things are recognized by their colour - of a stimulus by taking the physical properties of the stimulus and its surroundings into account [24]. It provides equations and methodologies for transforming physically measurable quantities to and from viewing condition specific perceptual attribute correlates [1, 12]. Within this context and as unrelated self-luminous stimuli have a dark surround, the new CAM needs to provide equations for transforming the physical properties of these stimuli to the traditional CAM correlates brightness, colourfulness, saturation and hue. Although some insights about colour perception can be gained during the development of this model, the doctoral research project is not intended to investigate physiological and psychological aspects of colour appearance, such as the effect of colour on, for example, emotion and well-being. Neither will the effect of, among others, context, self-luminous backgrounds and colour contrast be investigated. Only photopic, non-glare-inducing unrelated stimuli are investigated.

1.4.2 Applications of a CAM for unrelated self-luminous stimuli

The ability of predicting the colour appearance of unrelated self-luminous stimuli is useful for all conditions in which self-luminous stimuli are seen in dark surround conditions. An example is the increased use of LED signs in public areas for advertising purposes, possibly distracting drivers (see Fig.1.2) and contributing to light pollution. A CAM for unrelated self-luminous stimuli can be used to assist in the development of requirements for these LED signs, but also in the standardization of the appearance of marine, aviation or traffic lights viewed during the night.


Minister van Mobiliteit Ben Weyts (N-VA) laat effect op ons rijgedrag meten

Bedoeling: aandacht trekken
Resultaat: aandacht afgeleid

Er moet een einde komen aan de wildgroei van lichtgevende reclameborden langs de Vlaamse wegen. Dat zegt Vlaams minister van Mobiliteit Ben Weyts (N-VA). De borden worden overal neergepoot, vaak zonder vergunning. En vooral: zonder te beseffen hoe sterk ze bestuurders kunnen afleiden. Internationale studies wijzen al op een gevaar, maar minister Weyts laat nu zelf onderzoeken hoe verkeersonveilig ze zijn.

'Het is verboden om op de openbare weg reclameborden aan te brengen die bestuurders verblinden of in dwaling brengen', zo staat in de wegcode. Toch staan langs Vlaamse wegen talloze reclameborden die met bewegende lichtreclame en slogans zo veel mogelijk proberen op te vallen. 'Ik laat onderzoeken hoe verkeersonveilig de borden zijn', zegt bevoegd minister Ben Weyts (N-VA), in reactie op een vraag van Vlaams parlementslid Dirk De Kort (CD&V). 'Er gaat geen dag voorbij of er gebeuren zware ongevallen op onze wegen. Vorig jaar vielen er 400 doden. Bestuurders kunnen zich geen afleiding permitteren. Als minister kan ik automobilisten niet vragen om meer aandacht te besteden aan de weg, en tegelijk toestaan dat die aandacht steeds wordt afgeleid.'

Simulator

Het Agentschap Wegen en Verkeer (AWV) voert het onderzoek uit en zal onder andere in een simulator testen wat de lichtgevende en bewegende publiciteitsborden met een bestuurder doen. 'We onderzoeken welke parameters een invloed kunnen hebben. Zo kunnen de lichtsterkte en de snelheid waarmee boodschappen worden afgewisseld, mogelijk voor afleiding zorgen. We verwachten dit najaar de eindresultaten', zegt woordvoerster Ilse Luypaerts. Die resultaten zullen bepalen of er strengere criteria nodig zijn voor het plaatsen van reclameborden.

Een Nederlands onderzoek van de Stichting Wetenschappelijk Onderzoek Verkeersveiligheid wees in ieder geval al uit dat het 'heel lastig is om verlichte, knipperende en bewegende objecten op reclameborden te negeren'. De borden kunnen ook verwarrend zijn als ze vlak bij elektronische verkeersborden staan.

Bovendien worden de reclameborden zomaar overal neergepoot, vaak zonder stedenbouwkundige vergunning. 'We moeten daarom de wildgroei stoppen', zegt Weyts.

Fig.1.2. Article in 'Het Nieuwsblad' on 10/04/2015 about the distraction of bright self-luminous billboards.

In addition, the steps used in the new model for unrelated self-luminous stimuli can also be used to improve current CAMs for other appearance modes and thus also improve the quality of its applications. Furthermore, the results of the research could also be used in future studies of colour appearance and the research itself fits in the overall strategy of luminance -

or even appearance - based design in architecture. Luminance based design is a concept that allows to describe the technical performance of a lighting installation based on luminance as the basic criterion in contrast to illuminance [25, 26]. It allows lighting designers to design a room by taking into account luminance and luminance contrast of the entire visual field. Such a luminance based design has the possibility to improve visual comfort while reducing electricity consumption and avoiding glare.

Chapter 2

COLOUR APPEARANCE MODELLING OF UNRELATED COLOURS

As mentioned in the introduction, the perception of unrelated stimuli can be considered as being relatively simple and elementary compared to related stimuli. In this chapter, the main features of two existing CAMs for unrelated stimuli, CAM97u [16] and CAMFu [17], are discussed. It will become clear that colour appearance modelling of these stimuli is not quite simple at all. However, the complexity of CAM97u and CAMFu is not only a result of the complicated perception of unrelated stimuli, also the use of unclearly defined parameters and untested colour appearance phenomena comes into play…

2.1 CAM97u

2.1.1 Introduction

CAM97u is a CAM for unrelated colours developed by Hunt [15, 27]. An extensive description of the model was presented in the book *Measuring Colour* [16]. The model can be seen as a revision of Hunt's CAMs for related colours and is restricted to stimuli with a field of view (FOV) of 2°. Unfortunately, as mentioned before, the CAM97u model was not tested due to the lack of data [22]. Only recently, Fu performed visual experiments with unrelated stimuli and discussed the performance of the CAM97u model [28].

2.1.2 Input data

CAM97u uses the CIE 1931 *xy* chromaticity coordinates and photopic and scotopic luminance of the stimulus, the adapting field and the conditioning field as input:

- Stimulus: the colour element considered, a uniform patch with a FOV of 2°.
- Adapting field: the total environment of the stimulus.
- Conditioning field: the field seen just prior to viewing the unrelated colour stimulus. If there is no conditioning field, the chromaticity coordinates and photopic and scotopic luminance of this field are taken identical to the adapting field.

Even in a completely dark field, it is - according to Hunt [16] - not realistic to take the luminance of the adapting field as zero, because the stimulus, and scattering light from it in the eye, will provide an effective luminance of the adapting field above zero. The chromaticity of the adapting field is taken to be that of the equi-energy stimulus, EEW, because this is similar to the stimulus that appears most neutral to the dark-adapted eye [16, 29]. The photopic, L_A, and scotopic, L_{AS}, luminance of the adapting field are calculated from the photopic, L, and scotopic, L_S, luminance of the stimulus:

$$L_A = L^{2/3} / 200 \tag{2.1}$$

$$L_{AS} / 2.26 = (L_S / 2.26)^{2/3} / 200 \tag{2.2}$$

The photopic and scotopic luminance of a stimulus can be calculated from its spectral power distribution using the $V(\lambda)$ and $V'(\lambda)$ function, respectively.

These functions represent the average spectral sensitivity of the human visual perception for photopic and scotopic vision. Note that, if the spectral power distribution of the stimulus is not known, CAM97u offers an approximation to the scotopic luminance, L_S, based on the photopic luminance, L:

$$L_S / 2.26 = L \times (T / 4000 - 0.4)^{1/3} \tag{2.3}$$

Where T is the correlated colour temperature, if the stimulus has a chromaticity not too far from the Planckian locus (locus of points in a chromaticity diagram that represents chromaticities of the radiation of Planckian radiators at different temperatures).

2.1.3 Cone excitations

The first step of the model is to calculate the absolute XYZ tristimulus values from the luminance and chromaticity coordinates, both for the stimulus and for the conditioning field:

$$X = xL / y \qquad Y = L \qquad Z = (1 - x - y)L / y \tag{2.4}$$

From these tristimulus values, the cone excitations ρ, γ, β can be calculated. These cone excitations can be considered as the amounts of radiation usefully absorbed per unit area in the retina by the three different types of cones. It is desirable that these cone excitations are equal for the stimulus appearing most neutral to the dark-adapted eye [29]. Therefore, the equi-energy stimulus, EEW, has been used for normalisation. The cone excitations ρ, γ, β are calculated using the EEW-normalised Hunt-Pointer-Estevez matrix:

$$\begin{bmatrix} \rho \\ \gamma \\ \beta \end{bmatrix} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2.5}$$

2.1.4 Compressed cone responses

Non-linear compression

In the CAM97u model, the cone excitations ρ, γ, β are compressed into cone responses ρ_a, γ_a, β_a using a sigmoidal relationship that models the non-linear behaviour of visual responses (see Fig.2.1):

$$g_n(x) = f_n(x) + 1 = 40\left[x^{0.73} / (x^{0.73} + 2)\right] + 1 \tag{2.6}$$

The function, $g_{\mathrm{n}}(x)$, is based on the cone response compression of a rhesus monkey eye, investigated by Valeton and Norren [30]. It compresses the dynamic range of the responses, x, that have to be transmitted from the retina to the brain and provides a minimum and maximum level for the response. The term 1 in $g_{\mathrm{n}}(x)$ represents the noise of the signal. The minimum and maximum level models the threshold behaviour at low levels (value of 1) and saturation behaviour at high levels (value of 41). Such a compression is required to model the visual system over a large range in luminance levels [24]. In the central part of the function's operation range, the sigmoidal function performs similarly to a square root function (see Fig.2.1).

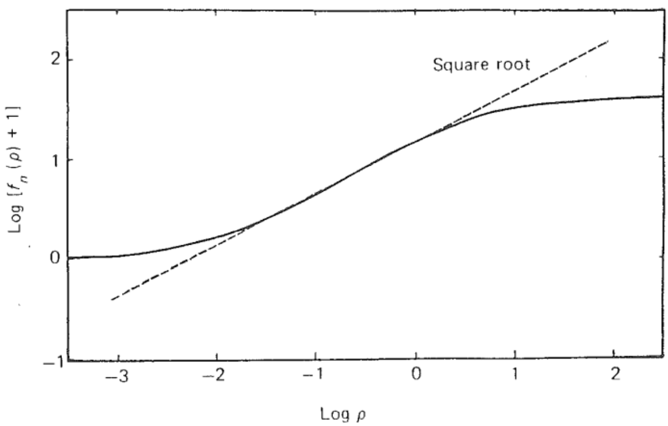


Fig.2.1. Sigmoidal function used to compress the cone excitations to cone responses. The log of the function, $\log[f_{\mathrm{n}}(\rho)+1]$, is plotted against the log of the radiation usefully absorbed, $\log \rho$ [16].

Adaptation

Prior to calculating the compressed cone responses, using Eq.2.6, the cone excitations are first modulated to account for the state of adaptation. In Hunt's CAM for *related colours*, a luminance adaptation and an approximate compensation for the effects of changes in the level and colour of the illumination are provided:

- The luminance adaptation is incorporated by a luminance adaptation factor, F_{L}, accounting for light adaptation over a wide range in luminance levels. Light adaptation is described by the Stevens and Hunt effect. According to the Stevens effect, a progressive reduction of the brightness of light colours and increase in the brightness of dark colours takes place as the level of illumination falls [31]. According to the Hunt effect, a progressive reduction of the colourfulness takes place as the level of illumination falls [31].

- The compensation for the effects of changes in the level and colour of the illumination, known as chromatic adaptation, is implemented by the use of chromatic adaptation factors, F_ρ, F_γ and F_β. According to the CIE definition, chromatic adaptation is a visual process whereby approximate compensation is made for changes in the colours of stimuli, especially in the case of changes in illuminants [1]. The factors F_ρ, F_γ and F_β are designed such that the equi-energy stimulus EEW always appears achromatic according to the model. They also take into account the effects that chromatic adaptation becomes less and less complete as the purity of the colour of the adapting light increases and becomes more and more complete as the luminance of the adapting light increases [2].

In his model for *unrelated colours*, CAM97u, Hunt still provides an adaptation very analogous to the one used in his model for *related colours* [15]. The luminance adaptation factor F_{L} was adopted from the CAM for related colours, yielding:

$$\begin{gathered} F_{\mathrm{L}} = 0.2k^4(5L_{\mathrm{A}}) + 0.1\left(1-k^4\right)^2\left(5L_{\mathrm{A}}\right)^{1/3} \\ k = 1/(5L_{\mathrm{A}}+1) \end{gathered} \tag{2.7}$$

The calculation of the chromatic adaptation factors F_ρ, F_γ and F_β was slightly modified by using the cone excitations of the conditioning field instead of the cone excitations of a reference white. For example, the chromatic adaptation factor for ρ is calculated as:

$$\begin{gathered} F_\rho = \left(1 + L_{\mathrm{A}}{}^{1/3} + h_\rho\right) / \left(1 + L_{\mathrm{A}}{}^{1/3} + 1/h_\rho\right) \\ h_\rho = 3\rho_{\mathrm{C}} / \left(\rho_{\mathrm{C}} + \gamma_{\mathrm{C}} + \beta_{\mathrm{C}}\right) \end{gathered} \tag{2.8}$$

When there is no conditioning field, these chromatic adaptation factors are set equal to 1. Additionally, in an attempt to take the effect of stimulus intensity on pupil diameter into account, the sensitivity is further adjusted by dividing each cone signal by a weighted summation of the three cone excitations, W:

$$W = \left[(1/3)(\rho + \gamma + \beta)\right]^{1/2} \tag{2.9}$$

In addition to this sensitivity adjustment, a factor $(L_A/L_C)^{0.2}$ was included to reduce the cone responses when L_C (luminance of conditioning field) is greater than L_A (luminance of adapting field). Finally, to provide an upper limit to the cone responses at very high levels of illumination, cone bleach factors $B_{\rho u}$, $B_{\gamma u}$ and $B_{\beta u}$ were introduced. For example, the cone bleach factor for ρ is calculated as:

$$B_{\rho u} = 10^7 / \left[10^7 + (5L_A)3\rho_C / \left(\rho_C + \gamma_C + \beta_C\right)\right] \tag{2.10}$$

All of this results in the following compression from cone excitations to cone responses:

$$\begin{aligned} \rho_a &= B_{\rho u}\left\{f_n\left[F_L F_\rho (L_A / L_C)^{0.2} \rho / W\right]\right\} + 1 \\ \gamma_a &= B_{\gamma u}\left\{f_n\left[F_L F_\gamma (L_A / L_C)^{0.2} \gamma / W\right]\right\} + 1 \\ \beta_a &= B_{\beta u}\left\{f_n\left[F_L F_\beta (L_A / L_C)^{0.2} \beta / W\right]\right\} + 1 \end{aligned} \tag{2.11}$$

2.1.5 Neural signals

The opponent colour theory suggests that there are three opponent neural channels: red versus green, blue versus yellow, and black versus white. According to the Natural Colour System (NCS), a perceptual colour model based on the opponent colour theory, the colours of an opponent channel are never perceived together: there is no “reddish green” or “yellowish blue”. The three neural signals used in CAM97u - an achromatic, a redness-greenness and a yellowness-blueness signal - are respectively obtained from a summation of a photopic (A_a) and scotopic (A_s) achromatic signal and two comparisons of the three cone differences signals (C_1, C_2 and C_3):

$$\begin{gathered} A_a = 2\rho_a + \gamma_a + (1/20)\beta_a - 3.05 + 1 \\ A_s = 3.05 B_{Su}\left\{f_n\left[F_{LS}\left(L_{AS} / L_{CS}\right)^{0.2}\left(L_S / 2.26\right)^{1/2}\right]\right\} + 0.3 \\ A = A_a + A_s - 0.26 \\ C_1 = \rho_a - \gamma_a \\ C_2 = \gamma_a - \beta_a \\ C_3 = \beta_a - \rho_a \end{gathered} \tag{2.12}$$

The achromatic signal, A, caries information mainly related to the intensity of the stimulus, reasonably assumed to be a weighted summation of the rod and cones responses. The weights in the summation of the cone responses, yielding the photopic achromatic signal, A_a, were chosen in accordance with the estimated relative number of cones of each cone type; ρ:γ:β is about 2:1:1/20. The scotopic achromatic response, A_s, is included to incorporate the rod response after adaptation. It is essentially a sigmoidal compression, adjusted with a rod bleach factor (B_{Su}), of the adapted scotopic luminance of the stimulus ($L_S/2.26$), whereby the adaptation depends on the scotopic luminance adaptation level (F_{LS}) and the compressed ratio of the adapting field luminance (L_{AS}) and the conditioning field luminance (L_{CS}) [16]. The calculation of the rod bleach factor B_{Su} and the scotopic luminance adaptation level F_{LS} is quite similar, but even more complicated, to the calculation of the cone bleach factors (Eq.2.10) and the luminance adaptation factor (Eq.2.7).

The cone difference signals (C_1, C_2 and C_3) carry information related to the colour, i.e. hue and colourfulness. In Hunt's model, the cone difference signals are equal to zero for achromatic stimuli and have a constant ratio C_1:C_2:C_3 for equal hues. By comparing the unique hue loci of the NCS with the curvatures of constant C_1/C_2 values for a set of cone spectral sensitivities based on the Stiles and Burch 1959 colour matching functions (CMFs) [32] normalized to the CIE Standard Illuminant C [33], Hunt developed a criterion for each unique hue. This comparison was performed in a transformed r'-g', $2b'/3$ chromaticity diagram (where r', g', b' are chromaticities corresponding to responses $R' = 1.05R$, $G' = 1.35G$, $B' = 0.60B$, with $r' = R'/(R'+G'+B')$) superimposed on the u', v' diagram (see Fig.2.2). A unique hue is defined as a hue that cannot be further described by the use of hue names other than its own, there are four unique hues: red, green, blue and yellow [1]. Simplified cone difference signals were used, based on a square root compression instead of a sigmoidal compression of the cone excitations. For example, the curve for which $C_1/C_2 = 1$ (or $\rho^{1/2} - \gamma^{1/2} = \gamma^{1/2} - \beta^{1/2}$) for the Stiles and Burch 1959, S_C normalized cone spectral sensitivities lies close to the NCS red locus and thus $C_1/C_2 = 1$ was chosen by Hunt as criterion for unique red. Also for the other unique hues, a criterion was developed by Hunt: $C_1 = C_3$ for unique green, $C_1 = C_2/11$ for unique yellow and $C_1 = C_2/4$ for unique blue by comparing the curvatures for different constant ratios of C_1/C_2 with the unique hue loci of the NCS scheme. More information of these criteria can be found in [3].

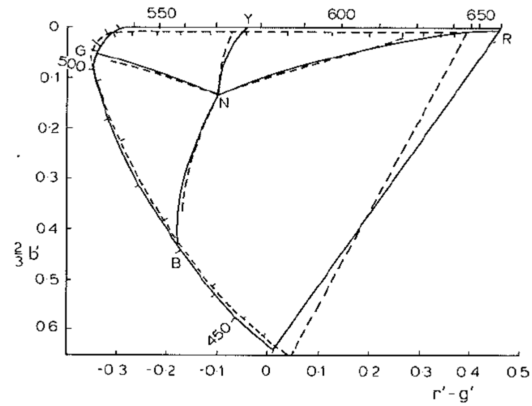


Fig.2.2. Comparison of unique-hue loci: (—) predictions corresponding to $C_1=C_2$ (curve R), $C_3=C_1$ (curve G), $C_1=C_2/11$ (curve Y), and $C_1=C_2/4$ (curve B) plotted in a transformed r'-g', $2b'/3$ chromaticity diagram; (- - -) unique red, green, yellow and blue loci of the NCS scheme plotted in the u', v' diagram [3].

According to the opponent colour theory, a yellow-blue and red-green opponent response in colour vision is assumed. Based on Hunt's unique red criterion, $C_1/C_2 = 1$ or $C_1=C_2$, the yellowness or blueness of reddish colours might be expected to be correlated with the extent to which C_1 is not equal to C_2. An approximate correlate of the yellow-blue opponent response, also called the yellowness-blueness b, could then be C_2-C_1. Similarly, due to the unique green criterion $C_3=C_1$, the correlate C_1-C_3 could represent the yellowness or blueness of greenish colours. The average of both correlates, $1/2(C_2\text{-}C_1+C_1\text{-}C_3)=1/2(C_2\text{-}C_3)$, could thus represent the yellowness-blueness b. Similarly, for the red-green opponent response, a correlate can be calculated based on the redness or greenness of yellowish ($C_1–C_2/11$) and bluish ($C_1–C_2/4$) colours. Instead of taking the average of these two correlates, the correlate for the redness or greenness of yellowish colours was taken as red-green opponent response, also called the redness-greenness a: $C_1–C_2/11$. Hunt made this choice based on experiments suggesting that unique yellow colours are more sharply defined than unique blue colours [34].

Because of the numerical distribution of the β cones (the least populous cone type involved in the yellowness-blueness) compared to the γ cones (the least populous principal cone involved in the redness-greenness), about 1/20, the yellowness-blueness in CAM97u is divided by $(1/20)^{1/2}$, that is by about 1/4.5. This is implemented by Hunt [3] in accordance with experimental results suggesting that, concerning the perceptibility of colour differences, a given change in retinal response has much greater effect in the red and green directions than in the blue and yellow [35]. Finally, the redness-greenness a and yellowness-blueness b is then indicated by:

$$\begin{gathered} a = C_1 - \left(C_2 / 11\right) \\ b = 1/2 \times \left(1/20\right)^{1/2}\left(C_2 - C_3\right) = 1/9\left(C_2 - C_3\right) \end{gathered} \tag{2.13}$$

Inconsistencies in the model:

The noise factor of the achromatic signal A described in Eq.2.12 (-0.26) is different compared to the one described in *Measuring Colour* (-2.31) [16], the latter being incorrect.

One of the first steps in CAM97u, is a transformation from XYZ tristimulus values (calculated using the CIE 1931 CMFs) to cone excitations ρ, γ, β using the EEW normalised Hunt-Pointer-Estevez matrix (Eq.2.5). However, the unique hue loci used in the model are obtained by using cone spectral sensitivities based on the Stiles and Burch 1959 CMFs [32] normalized to CIE illuminant C.

2.1.6 Hue correlate

As in other models, the hue angle in CAM97u is calculated as the arctangent of the yellow-blue opponent dimensions divided by the red-green opponent dimensions:

$$h = \tan^{-1}\left(b / a\right) = \tan^{-1}\left[\frac{1/9(C_2 - C_3)}{C_1 - (C_2/11)}\right] \tag{2.14}$$

When the hue is expressed in terms of the proportions of the unique hues perceived to be present in the stimulus a more perceptual meaningful attribute, hue quadrature H, is obtained. Hue quadrature H expresses the hue in a number between 0 and 400, where red is represented by 0 (or 400), yellow by 100, green by 200 and blue by 300. A hue quadrature of 240 represents a particular teal stimulus containing 60% green and 40% blue. The difference between hue angle h (Eq.2.14) and hue quadrature H, is illustrated in Fig.2.3. In the hue plot (left), the value of h is the angle between a horizontal line drawn from the origin towards the right and the line joining the origin to the point representing the stimulus colour. The positions of the unique hue lines are shown in this diagram by the coloured full lines, R, Y, G, and B. The broken lines represent the angular positions where the perceptually midway of the adjacent pairs of unique hues lie, e.g. 50% red and 50% yellow. These broken lines are not at equal angular spacing in the hue angle. In order to obtain a hue quadrature, the positions of the unique hue lines should be orthogonal, with the opponent colours in opposition, and the perceptual midway point spaced at regular intervals (see Fig.2.3 right).

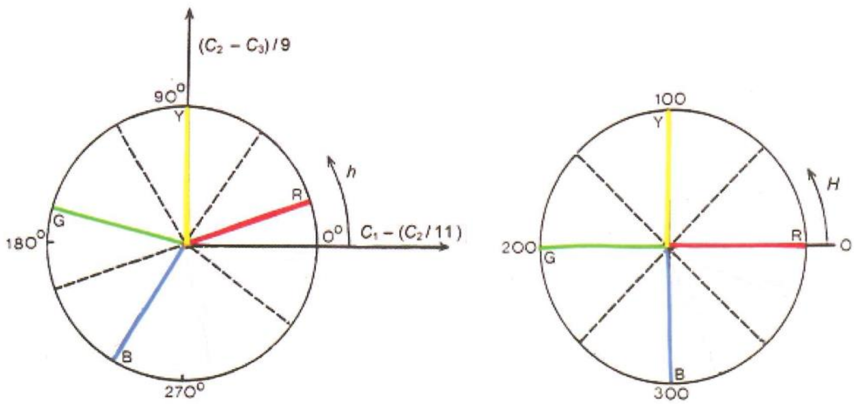


Fig.2.3. Left: hue angle, h, shown in a plot of yellowness-blueness, b, against redness-greenness, a. Right: hue quadrature, H, shown in a plot where unique red and green are opposite one another, and unique yellow and blue are also opposite one another and at right-angles to red-green directions [31].

The calculation of hue quadrature is given by:

$$H = H_i + \frac{100(h' - h_i)/e_i}{(h' - h_i)/e_i + (h_{i+1} - h')/e_{i+1}} \tag{2.15}$$

With h_i the unique hue angle, H_i the unique hue quadrature, and $h'=h+360$ if $h<h_1$, otherwise $h'=h$, a value i chosen so that $h_i \le h' < h_{i+1}$ and e_i the eccentricity factor of the unique hues (see Table 2.1).

Table 2.1. Overview of the unique hue data used in CAM97u for calculating the hue quadrature *H*.

Unique hue	Red	Yellow	Green	Blue	Red
i	1	2	3	4	5
h	20.14	90.00	164.25	237.53	380.14
e_i	0.8	$0.7*L/(L+10)$ $+0.3*10/(L+10)$	1.0	$1.2*L/(L+10)$ $+0.2*10/(L+10)$	0.8
H_i	0.0	100.0	200.0	300.0	400.0

The eccentricity factor e_i was introduced to compensate for the differences in strength of perceptual colorization that occur around the hue circle. For example, as indicated by CIE $s_{u',v'}$, the saturation of a yellow stimulus can never be as high as that of blue stimulus. So the eccentricity factors can be regarded as the weights with which the responses for the hues contribute toward saturation [3]. The eccentricity values were experimentally obtained in a cone signal space, resulting in about 0.65 for red, 0.5 for yellow, 1.0 for green and 1.45 for blue. In the cone response space used in CAM97u, where the central part of the sigmoidal compression function can be approximated by a square root, these eccentricities have respective values of 0.8, 0.7, 1.0 and 1.2. For a more thorough discussion on eccentricity the reader is referred to [3].

The unique hue angles listed in Table 2.1 were obtained using the hue equation, Eq.2.14, and Hunt's criteria for unique hues (see above). For example, using the criterion for unique red, $C_1 = C_2$ or $\rho^{1/2} - \gamma^{1/2} = \gamma^{1/2} - \beta^{1/2}$, the unique red hue angle h_1 can be calculated:

$$\begin{aligned} h_1 &= \tan^{-1}(\mathrm{b}/\mathrm{a}) = \tan^{-1}\left[\frac{1/9(\rho_\mathrm{a} + \gamma_\mathrm{a} - 2\beta_\mathrm{a})}{\rho_\mathrm{a} - 12\gamma_\mathrm{a}/11 + \beta_\mathrm{a}/11}\right] \\ &= \tan^{-1}\left[\frac{1/9(3\gamma_\mathrm{a} - 3\beta_\mathrm{a})}{(10\gamma_\mathrm{a}/11 - 10\beta_\mathrm{a}/11)}\right] = \tan^{-1}(33/90) = 20.14° \end{aligned} \tag{2.16}$$

This was also done for the other unique hues, yielding 90° for yellow, 164.25° for green and 237.53° for blue.

Finally, a rough estimate of the Bezold-Brücke effect was included by making the eccentricity factors of yellow and blue dependent on the luminance L of the stimulus [15, 16]. The Bezold-Brücke effect states that the apparent hue of a colour is affected by the luminance level of the stimulus [27].

2.1.7 Saturation and colourfulness

Saturation, defined by the CIE as colourfulness judged in proportion to its brightness, is defined in CAM97u as a combination of redness-greenness a and yellowness-blueness b [3]. Hunt found that the simplest combination giving the same smooth contours for constant saturation in a chromaticity diagram, as found in practice, was a square root of the sum of squares of a and b [3]. Next, several factors were added. The *first* one is the eccentricity factor e, already described above. To calculate the eccentricity factor for hues other than the unique hues a linear interpolation is used. A *second factor* is the chromatic surround induction factor N_c, which makes allowance for the fact that dark or dim surrounds can reduce the colourfulness of the stimulus. Because unrelated colours have per definition a dark surround the N_c factor is set to a fixed value of 0.5. Saturation is then calculated by dividing the product of the root-sum-square of a and b and the two corrective factors by a correlate of brightness. In the model $\rho_a + \gamma_a + (21/20)\beta_a$ is used, with the factor 21/20 included to facilitate reversing the model and to account for the physiologically plausible combined strengths of the achromatic and colour difference signals [15]. *Finally*, some constant factors were added to allow for cross-channel noise between the cone responses, 10/13, and to give convenient numbers, 50 and 100. A correlate of saturation s, is then given by

$$s = 50(a^2 + b_{tu}^2)^{0.5} 100e(10/13)N_c / [\rho_a + \gamma_a + (21/20)\beta_a] \qquad (2.17)$$

With b_{tu} equal to $bL/(L+0.1)$ to include the luminance tritanopia factor for unrelated colours. This luminance tritanopia factor predicts a loss of yellowness-blueness at progressively lower and lower luminance levels [2]. In other words, observers tend to become more and more tritanopic, i.e. yellow-blue deficient, as the luminance decreases.

Next, the colourfulness is calculated by multiplying the saturation with a correlate of brightness, being the luminance adaptation factor F_L:

$$M = sF_L^{0.15} \qquad (2.18)$$

Inconsistency in the model:

Because saturation is defined by the CIE as the colourfulness of a stimulus judged in proportion to its brightness, the colourfulness should be calculated prior to the saturation. In addition, a simple formula using the actual brightness attribute would have been a more logical choice instead of a brightness correlate such as the luminance adaptation factor F_L in Eq.2.18. Finally, by using different brightness correlates at different stages (e.g. saturation, colourfulness) the model lacks internal consistency.

2.1.8 Brightness correlate

Finally, the brightness Q is calculated as a weighted combination of the achromatic signal and the colourfulness. This combination is based on the idea that while the achromatic signal is the main contribution to the perception of brightness, colourfulness is known to have an effect as well (Helmholtz-Kohlrausch effect):

$$Q = \left[1.1\left(A + M / 100\right)\right]^{0.9} \tag{2.19}$$

This formula is partially based on the work of Bartleson [36] where, amongst others, the brightness of unrelated colours under dark adapted conditions was investigated. Although Bartleson predicted the brightness using the cube root of the luminance, the prediction above (Eq.2.19) is built in a square root response space. Hunt also never validated his proposed functional relationship between brightness and the achromatic signal and colourfulness: the factor 1/100, chosen arbitrarily, will be shown to be severely underestimating the Helmholtz-Kohlrausch (H-K effect) for unrelated self-luminous colours.

2.2 CAMFu

2.2.1 Introduction

CAMFu is a CAM for unrelated colours developed by Fu et al. [17, 37], and highly inspired by CAM97u and CIECAM02. It is a combination of these two models, based on the performance of the models for visual data obtained in a magnitude estimation experiment for unrelated colours under photopic and mesopic conditions. In the experiment, 10 observers evaluated the perception of brightness, colourfulness and hue for different sized, unrelated stimuli with a luminance between 0.013 and 60 cd/m². For photopic viewing conditions, stimuli with a FOV of 0.5° and 10° were used.

2.2.2 Input data

The input data of CAMFu are the CIE 1931 xy chromaticity coordinates and the photopic and scotopic luminance of the stimulus. Furthermore, the adapting luminance L_A, a luminance factor of the background Y_b, and the stimulus size, in degrees, is needed. The adapting luminance is taken as 20% of the luminance of a white stimulus in the group of stimuli being considered. Therefore, the luminance factor of the background is fixed to a value of 20.

For some steps, an adopted white and some surround parameters are required (cfr. the equations of CIECAM02). In CAMFu, the adopted white is taken to be that of the equi-energy stimulus, EEW, with a luminance of 100 cd/m², and the surround parameters are set to those under dark viewing conditions of CIECAM02: $c = 0.525$, $N_c = 0.8$ and $F = 0.8$.

2.2.3 Cone excitations

Similar to the procedure followed in CIECAM02, however without the inclusion of a chromatic adaptation transform, the cones excitations ρ, γ, β are calculated from the *relative* CIE 1931 *XYZ* tristimulus values using the Hunt-Pointer-Estevez matrix (Eq.2.5). The relative *XYZ* tristimulus values are, in turn, calculated from chromaticity coordinates *xy*.

2.2.4 Compressed cone responses

The compression from cone excitations to cone responses used in CAMFu is the same as the one in CIECAM02, where again a sigmoidal curve is used. Compared to CAM97u (and CIECAM97s), in CIECAM02 a modified sigmoidal curve is used as it has been shown that the change in saturation due to a change in the adapting luminance, is much larger in the CIECAM97s prediction than likely to occur in practice [38]. Changing the power in Eq.2.6 to 0.42 (instead of 0.73 as in CAM97u and CIECAM97s) effectively alters the square root relationship in the central part of the sigmoidal curve to a cube root, improving the saturation prediction.

As mentioned earlier, the actual response of the cones is not only dependent on the intensity of the stimulus, but also on the state of adaptation. Similar to both CAM97u and CIECAM02, CAMFu implements a luminance adaptation factor F_L when calculating the cone responses. For example, the compressed cone response ρ_a is calculated as:

$$\rho_a = \left[400\left(F_L\rho / 100\right)^{0.42}\right] \Big/ \left[\left(F_L\rho / 100\right)^{0.42} + 27.13\right] + 0.1 \qquad (2.20)$$

2.2.5 Neural signals

The transformation from compressed cone responses to neural signals - i.e. the calculation of the cone difference signals C_1, C_2 and C_3 (Eq.2.12), the criteria for achromacy and constant hue, and the equations for the neural signals a and b (Eq.2.13) - is largely identical to the one used in CAM97u, with the exception of the derivation of the achromatic signal A. The latter is calculated as a weighted summation of the photopic, A_a, and scotopic, A_s, part of the signal:

$$A = A_a + kA_s \tag{2.21}$$

With k a constant obtained empirically using visual data and used to determine the ratio between cone and rod contributions, respectively equal to $-5.3 \times \log_{10}(L) + 44.5$ and $-5.9 \times \log_{10}(L) + 50.3$ for $0.5°$ and $10°$ stimuli. The scotopic part of the achromatic signal, A_s, is obtained by a compression of the scotopic luminance, $L_s^{0.42}$. The photopic achromatic signal is obtained from CIECAM02, which is calculated very similar as in CAM97u (Eq.2.12).

Inconsistencies in the model:

In CAM97u and CIECAM97s, the yellowness-blueness b was reduced by 9 because this signal should have less weight than the redness-greenness signal. 1/9 is approximately equal to 1/2 times $[1/20]^{1/2}$, with 1/2 resulting from taking the average of the colour correlates and $[1/20]^{1/2}$ obtained from the numerical distribution of the β cones compared to the γ cones. However, as in CIECAM02 and CAMFu the central part of the compression to cone responses follows a cube root relationship instead of a square root relationship, this reduction should be equal to 1/2 times $[1/20]^{1/3}$, approximately 1/5.42.

Furthermore, despite adopting a cube root response space, CAMFu keeps the original yellowness-blueness and redness-greenness equations of CAM97u, which are based on Hunt's unique hue criteria derived in a square root response space, resulting in further errors to the model.

2.2.6 Hue correlate

The hue angle h and hue quadrature H of CAMFu (and CIECAM02) are calculated similar to CAM97u (Eq.2.14 and 2.15). The only difference is that the Bezold-Brücke effect was not included, yielding unique hue eccentricities for yellow and blue of 0.7 and 1.2, respectively.

Inconsistencies in the model:

It is important to note that the prediction of the hue angle and hue quadrature has been influenced by the incorrectly copied redness-greenness a

and yellowness-blueness b signals from the CAM97u and CIECAM97s model. In addition, the calculation of the unique hue angles also depends on these signals. For example, the factor 1/9 in the calculation of the unique red hue angle (Eq.2.16) should be changed into 1/5.42, yielding to an unique red hue angle in CAMFu (and CIECAM02) of 31.34° instead of 20.14°. This incorrect factor of b has an influence on almost all CAMFu (and CIECAM02) colour attributes.

In addition, the eccentricity factors of CAMFu (and CIECAM02) were also incorrectly 'copied' from those of the CAM97u and CIECAM97s models. In the latter, these eccentricity factors 0.8, 0.7, 1.0 and 1.2, for unique red, yellow, green and blue, respectively, were obtained based on the square root of experimental data (see above). However for CAMFu (and CIECAM02), a cube root relationship should have been used instead, yielding 0.87, 0.79, 1.00 and 1.13, respectively.

2.2.7 Colourfulness

In contrast to CAM97u, colourfulness is determined in CAMFu prior to the calculation of the saturation. Although the correlate of colourfulness is based on CIECAM02, it is slightly modified by a scaling factor, K_{M}, to account for the size of the stimulus, 0.5° or 10°. Because some factors are kept constant in CAMFu, the calculation of colourfulness can be simplified to:

$$M = 0.90 K_{\mathrm{M}} \left\{ \frac{\left[50\left(a^2 + b^2\right)^{0.5} 100e\left(10/13\right) N_{\mathrm{c}} \right]}{\left[\rho_{\mathrm{a}} + \gamma_{\mathrm{a}} + (21/20)\beta_{\mathrm{a}}\right]} \right\}^{0.9} F_{\mathrm{L}}^{0.25} \tag{2.22}$$

With K_M respectively equal to 0.9 and 1 for 0.5° and 10° stimuli. This colourfulness is closely related to the 0.9$^{\text{th}}$ power of the colourfulness of CAM97u.

2.2.8 Brightness and saturation correlate

Quite similar to CAM97u, the brightness is determined by taking a weighted summation of the achromatic signal and the colourfulness, however without the 0.9 exponent:

$$Q = A + M / 100 \tag{2.23}$$

Finally, the saturation is calculated according to its CIE definition, i.e. the colourfulness M relative to the brightness Q [1], but introducing a square root compression:

$$s = 100(M / Q)^{0.5} \tag{2.24}$$

2.3 Concluding remarks

In the absence of an agreed model for unrelated stimuli, Hunt developed CAM97u. Instead of being based on experimental results using unrelated stimuli, it is developed using his former models for related colours and slightly modified based on some previous experimental results and known appearance phenomena of unrelated stimuli. CAM97u was not developed any further, nor tested. It took about 15 years before an alternative model was given by Fu: CAMFu. In addition to some modifications based on the widely used CIECAM02 model for related colours, Fu's model is very similar to CAM97u.

Both models are a complex collection of parameters trying to include as many as possible colour appearance phenomena. However, the model parameters were often chosen rather arbitrarily, by trial and error or semi-physiologically and often without taking visual data into account. Today, it is difficult to understand the background and motivation for the introduction of all these parameters while it becomes a real challenge to correctly use the models in a specific application. Probably, this explains the lack of existing applications making use of CAM97u. Furthermore, the missing background about the development of each step, makes it tough to extend the model. For example, changing one of the first steps in the model (using a cube root instead of a square response space) has an influence - sometimes untraceable - on almost all parameters in the subsequent steps of the model. This results in a lot of inconsistencies. In fact, a more extensive investigation of the models' performance is needed. Only by testing the predictive performance of the models with independent experimental results, clear conclusions about the usability of the models can be made. However, it might also be a good time to try to start anew and build a model based on the solid achievements of visual science and experimental evidence.

Chapter 3

EXPERIMENTAL SETUP AND METHODS

How would you describe the colour appearance of the stimulus in front of you?

"It's red."

Can you be a little bit more precise?

"It's very red."

Hmmm. Could you give other words, besides red, describing your perception of the stimulus?

"Oh, this is a difficult one. I would say …. it is the same red as a good red wine. Not the red of Coca-Cola, more winy red. Is that a better answer?"

For the investigation of the appearance of stimuli several activities are required. In addition to designing and creating an experimental setup, choosing the perfect stimuli to be evaluated and looking for observers, it is important to ask the right questions. As observers are usually naïve with respect to the quantitative evaluation of a coloured stimulus, before the main experiment, a learning experiment is necessary. This learning experiment lets them become familiar with the scaling techniques to ensure a reliable evaluation of their perception. In this chapter the basic tools to investigate colour perception are introduced and the experimental setup and methods used in our investigation of the colour appearance of unrelated self-luminous stimuli are described.

3.1 Experimental setup

Two experimental setups were used in this investigation: a darkened room with a LED module in a viewing box with a diffusor on top, and another darkened room with an LCD monitor.

3.1.1 LED module with diffusor

The largest part of the experiments investigating the colour appearance of unrelated self-luminous stimuli was performed using a diffused LED module in a specially designed viewing room (see Fig.3.1 (left)). The viewing room of 3.1 m wide by 5.8 m long by 3.5 m high with black walls, a grey ceiling and a greyish black floor carpet was created to generate the unrelated self-luminous stimuli for the experiment. In the center of one wall and surrounded by a dark surround, a circular stimulus with a diameter of 37 cm was created. Observers were seated at a distance of 211 cm to generate a stimulus field of view of approximately 10°. The stimulus is produced by a number of red, green, blue and white light-emitting diodes (LEDs) mounted inside a white cylindrical cavity covered with a diffusor. By controlling the drive current of each LED using a DMX digital communication network the colour and luminance of the stimulus can be changed (see Fig.3.1 (right)). A heat sink and active cooling ensured a sufficient stable and reproducible light output.

Fig.3.1. (left) Experimental setup of the LED module. (right) Example of a stimulus under dark viewing conditions.

For the experiments, coloured stimuli with a certain luminance and having a wide chromaticity gamut were carefully selected. The luminance level was always chosen in order to ensure photopic viewing conditions without any glare effect. All colorimetric and photometric quantities were determined from spectral measurements using a spectroradiometer (MS260i Oriel instruments spectrograph or QE65000 Ocean Optics) and a suitable calibration.

The luminance uniformity of the stimulus area was checked by measurements with a two-dimensional luminance camera (MURATest by Eldim). The luminance of the stimulus was found to gradually decrease (to approximately 20% around the mean) from the centre to the edge. Observers are not aware of this variation as the human eye is rather insensitive to low spatial frequencies [39]. The background, consisting of a black curtain, provided a 0.01 cd/m² adaptive field.

3.1.2 LCD monitor

An experiment validating the visual results obtained using the LED module setup and an experiment investigating the effect of stimulus size on the brightness perception was performed using an LCD monitor. In a viewing room of 3.1 m wide by 5.8 m long by 3.5 m high with black walls, ceiling and floor, a wide gamut LCD monitor (Eizo ColourEdge CG246, 24") was placed against one wall. On this monitor circular stimuli with a FOV of 1° to 30° were presented to observers seated in front of the monitor with a fixed chinrest (Fig.3.2). All colorimetric and photometric quantities of the stimuli and background were determined from spectral measurements using a spectroradiometer (QE65000 Ocean Optics) and a suitable calibration.

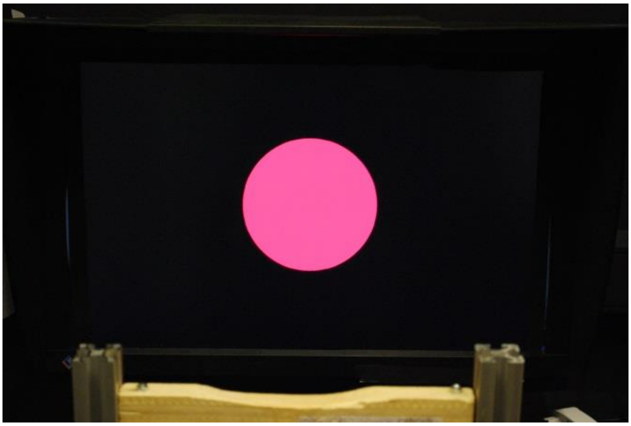


Fig.3.2. Experimental setup with a stimulus having a FOV of 20°.

The luminance uniformity of the stimulus area was again checked by measurements with the two-dimensional luminance camera (MURATest by Eldim). The luminance of the stimulus was found to be approximately constant over the stimulus area (maximum deviation of 3% around the mean). Again, observers were not aware of this variation. The background provided an adaptive field with luminance below 0.5 cd/m².

3.2 Investigating colour perception

The study of colour perception of unrelated self-luminous stimuli and its modelling in a CAM requires visual experiments in which observers are asked to evaluate absolute colour attributes. Several commonly used experimental methods for the collection of visual data of these absolute attributes are briefly discussed below. Some of them were also compared in a series of preliminary experiments.

3.2.1 Paired comparison method

In a paired comparison test of m stimuli, each of the $m(m\text{-}1)/2$ pairs is presented to r observers in one order, and to r observers in the other [40]. Mostly, the method is used for cases in which the observer's preferences are expressed on a 7 or 9-point scale. For example: "Rate the brightness of sample i on the right compared to that of sample j on the left, by use of the following preference scale; (3) i is much brighter than j; (2) i is brighter than j; (1) i is slightly brighter than j; (0) i and j are of equal brightness; (-1) i is slightly less bright than j; (-2) i is less bright than j; (-3) i is much less bright than j'. When observers rate both stimuli to be equally bright, they can still be asked to indicate a preference. As such, the scaling method can be combined with a two-alternative forced choice (2AFC) task [41]. To account for potential within-pair order effects, optimal order designs need to be applied [42], e.g. all stimuli must be compared with one another, every stimulus must appear as often on the left as on the right of a pair,…

Pro's

- Forced choice and interval scaling can be used together, extending the possible statistical analyses of the visual data.

Contra's

- Although incomplete experiments are possible, in an ideal paired comparison experiment all possible pairs need to be compared (taking a huge amount of time).
- Not all perceptual attributes can be evaluated: only brightness and colourfulness, but not hue.
- Visual tests to develop and investigate CAMs generally do not use the paired comparison method.

3.2.2 Magnitude estimation method

Magnitude estimation is a method in which observers are asked to quantify (e.g. numerically or graphically) their magnitude estimate of one or more perceptual attributes [43, 44]. For example: "Give a value to the brightness of sample *i* on the right compared to that of sample *j* on the left which is attributed a value of 50." These absolute perceptual values can be used directly to test various existing colour models or to develop a completely new CAM [9]. It is essential that each observer clearly understands the perceptual attributes being scaled and that the observers are familiar with this scaling method.

Pro's

- The method permits observers to give simple, absolute values to familiar colour attributes.
- The method can be used to judge all perceptual attributes of interest: brightness, colourfulness and hue.
- It is the most commonly used method to develop and investigate CAMs.
- All possible stimuli need to be evaluated by each observer only once, which substantially shortens the experiment duration.

Contra's

- Observers need to be trained to become familiar with the evaluation method.
- A reference stimulus - approximately in the middle of the perceptual range - needs to be chosen to anchor the perceptual attributes being scaled.

3.2.3 Matching method

The matching method is a two-interval adjustment task in which observers are presented with two stimuli in spatial or temporal juxtaposition, with one stimulus being the reference [45]. Observers are instructed to vary an attribute of the second stimulus (e.g. the brightness, by adjustment of its luminance) until it matches as closely as possible the attribute of the reference stimulus.

Pro's

- The method permits observers to evaluate attributes without asking them to give a value.
- The method can be used to judge brightness, colourfulness and hue, although it is only easily applicable for brightness.

Contra's

- Difficulty in avoiding starting position bias.
- Either a reference stimulus needs to be chosen or (preferably) all possible pairs are examined.
- No absolute values are given for the colour attributes, making it difficult to use the observer results when developing a CAM.
- Generally not used to develop and investigate CAMs.

3.2.4 Untested additional methods

Adjustment

Adjustment is a method in which observers are instructed to adjust an attribute of a stimulus to an optimum, preferred or ideal level [45]. Since a CAM predicts absolute values of the colour attributes and the adjustment method provides only physical data of the 'ideal' stimulus, the method has not been used.

Discrimination

Discrimination is a method in which observers are presented with two stimuli [45]. The observers are instructed to report which scene is, for example, brighter. This is usually a forced choice task. This method can be combined with the paired comparison method.

Ranking

Rank order methods require observers to rank stimuli in terms of an attribute, e.g. brightness [46]. The rank of stimuli can however also be obtained from the data of a magnitude estimation experiment, the preferred method in colour appearance modelling.

Rating

Rating is a method in which observers are instructed to use a rating scale to describe the appearance of a stimulus [45], e.g. a 4-point response scale representing a bright-dim axis with intervals labelled *very bright* (1), *bright* (2), *dim* (3) and *very dim* (4). The magnitude estimation method is kind of a rating method with a lot of intervals [9].

3.2.5 Additional considerations

Although several scaling methods can be used to evaluate a stimulus, only the magnitude estimation method permits observers to give simple, absolute values to familiar colour attributes. In addition, these absolute perceptual values can be used directly to test various existing colour models or to develop a completely new CAM [9]. Therefore, the visual experiments performed in this doctoral research project mainly followed the magnitude estimation approach. Findings of the preliminary experiments and suggestions found in literature were taken into account to design the experimental procedure.

Observers

According to the *ASTM International standard test method for unipolar magnitude estimation of sensory attributes* [43], a panel of 15 to 20 observers can produce data of adequate precision and reproducibility.

Typically, colour appearance models are developed for observers not suffering from any serious colour deficiency. Therefore, all observers should be tested for colour blindness, for example using the Ishihara 24 plate Test for Colour Blindness, the Farnsworth-Munsell 100 Hue Test [47] or other.

As the general validity of a model depends on the diversification of the observer panel, a balance should be found between naive outsiders and experienced insiders and between genders [45]. In addition, to obtain reliable data, observers should be dedicated to their task. Before each experiment in this doctoral research project, observers were explained why the research and their participation is important. Outsiders were rewarded with a voucher - approximately having a value of 10 euro per hour of presence - and all observers were pampered with candy and soda.

Observer variability

In each experiment, intra- and inter-observer variability should be assessed to respectively obtain the variance between repeated evaluations of the same observer and between the evaluation of different observers [9, 48]. By taking the mean of all intra-observer and of all inter-observer results, the observer variability can be compared to those of other experiments. Furthermore, the predictive performance of a model can be evaluated by comparing the inter-observer variance on one side with the variance between the average observer perception and the model prediction on the other side [17].

Fatigue

Preliminary experiments indicated that, to reduce the influence of fatigue, approximately after each half an hour of evaluating stimuli, a short break of 15 minutes should be taken.

Unrelated stimuli

To keep the stimuli unrelated, only a successive (instead of a simultaneous) presentation of the test and reference stimuli can be selected.

Evaluation

Depending on the method, the colour attributes of a stimulus can be evaluated using several 'measurements': explicit measurements - such as magnitude estimates of the brightness - and implicit measurements - such as brain activity, eye-movements, and behavioural measurements. Explicit measurements can be used directly to develop CAMs: by using a slide bar, writing down numbers, or giving numbers to the experimenter, observers rate their perception directly. For pure unrelated colours, only an oral evaluation can be used, as the use of a 'visible' slide bar or scale would make them effectively related colours.

Amount of white as a perceptual attribute

In a series of preliminary experiments, the colour terms attributed to the colour appearance of unrelated colours have been investigated using a panel of 10 naïve observers. Brightness (Dutch: "helderheid") and hue (Dutch: "kleurtint"), or synonyms for them ("donker - dark", "licht - light", "fel - intense or vivid", "opvallend - eye catching", "klaar - clear or light", "intens - intense",… for brightness and "groen - green", "cyaan - cyan", "rood - red",… for hue), were readily reported as perceptual attributes. The term colourfulness (Dutch: "kleurigheid") was never used. Words such as "bleek - pale", "licht - light", "afgetrokken van de zon - sun bleached", "bevat wit - contains white", "afgemat - matt", "dof - dull",… were given as alternatives for colourfulness. Note that most observers used the word "licht - light" twice, once for describing the colourfulness and once as referring to the brightness attribute. In another experiment, the scaling of the attributes was investigated and again no difficulties were found for brightness and hue. In contrast, observers seemed very unfamiliar with the term colourfulness and were unable to evaluate this attribute without a time consuming training. To increase the relevance of a CAM for practical applications, correlates which can be easily assessed by naïve observers should preferably be used. It turned out that instead of rating colourfulness, observers were more comfortable evaluating the "*amount of white versus non-white*" perceived to be present in a stimulus. Half of the (naïve) observers used the term 'amount of white' as alternative for colourfulness and additionally, after an experiment in which both colourfulness and "amount of white" were evaluated, observers assigned "amount of white" as the most easy to evaluate and the one that best reflected their perception. As a few observers also saw "grey" in some (darker) neutral stimuli, the amount of white was later extended to cover the concept "amount of neutral versus coloured".

Reference

When using the magnitude estimation method, a reference stimulus needs to be chosen to anchor the perceptual attributes being scaled [43]. For brightness, as a first estimate, a white, achromatic stimulus approximately in the middle of the luminance range of the test stimuli can be chosen. This reference can be presented repeatedly throughout the experimental session after one or several test stimuli. One can choose to give a fixed value to the brightness of this reference (e.g. a value of 50) or let the observer choose a value. Another possibility is to use the previously evaluated stimulus as reference (with the first stimulus given a particular value). From preliminary experiments, a fixed reference repeatedly presented after each test stimulus and having a brightness value of 50 was found to be the best approach.

For colourfulness, a saturated stimulus - with a particular value for its colourfulness - needs to be chosen as a reference. The hue of this reference is however a key issue: the colourfulness can be evaluated by each observer against a reference of the same hue as the test stimulus (e.g. a red reference for reddish stimuli,…) or against a grey stimulus using a greyscale. Another possibility would be that a group of the observers scales against one reference, another group against another reference (e.g. a fixed green reference for the first group of observers and a blue reference for the second group of observers,…). Each of these possibilities introduces problems: the lack of a single colourfulness scale for all hues, the unknown relationship between greyscale and colourfulness, the small amount of observers for each reference as the observers are split up into groups,…

For the 'amount of white' and hue, observers use an internal reference when evaluating the stimuli, thereby substantially simplifying the experiment.

Learning experiment

It is essential that each observer clearly understands the perceptual attributes being scaled and that the observers are familiar with the scaling method [43]. Therefore, in this doctoral research project naïve observers completed a straightforward exercise in which they were asked to rate the length of a line in comparison with a line of length 100, similar to a method described in the *ASTM International standard test method for unipolar magnitude estimation of sensory attributes* [43]. Before each experiment, observers also completed a learning experiment with a set of stimuli similar to the ones used in the experiment to help them to become familiar with the rating technique as applied to coloured stimuli and to make them aware of the colour and luminance range. Note that observers were not trained to give a 'forced' answer about their perception. The observers participated in the 'learning experiment' only to ensure they become aware of the colour range and its accompanying perceptual attribute range and to produce data of adequate precision and reproducibility [43].

Sequence

As preliminary experiments had shown observers to have difficulty rating all three attributes at once, the brightness should be rated separately from the hue and amount of white. In addition, the stimuli should be randomly arranged in two series, each one being evaluated by half of the observers to avoid possible bias due to the series sequence [45, 49]. Prior to each experiment, observers should also adapt to the dark viewing conditions.

3.3 Experimental procedure

Based on the suggestions discussed above, an experimental procedure was developed. Before each experiment, naïve observers participated in a learning experiment of 30 minutes to become familiar with the magnitude estimation method. They completed some straightforward exercises and, like the experienced observers, they completed an experiment with a similar set of stimuli as used in the main experiment to help them become familiar with the rating technique as applied to coloured stimuli and to make them aware of the colour range.

To reduce the influence of fatigue in the experiments, each experimental session was limited to about 30 minutes. A break of about 15 minutes was offered between each session. In most of these sessions, in addition to the test stimuli, 5 stimuli as 'warming up' and 10 stimuli used to calculate observer variability, were presented. The stimuli were always randomly arranged in two series, each one being evaluated by half of the observers to avoid possible biases due to the series sequence [49]. In experiments where all three attributes were to be evaluated, about half of the observers started with the scaling of brightness, while the other half started with the scaling of the amount of white and the hue.

Each stimulus was presented to the observers for 15 seconds. Between these stimuli, a reference achromatic stimulus was shown for 5 seconds. Each experiment session started by showing this reference achromatic stimulus. To keep the coloured stimuli unrelated, the reference stimulus was shown in temporal juxtaposition with the test stimulus. Although this approach can possibly induce small memory errors, it was preferred to the simultaneous presentation at adjacent spatial locations. Total darkness never occurred in order to reduce the possibility of temporary blindness and afterimages. Before the experiment, the observers adapted to the dark viewing conditions for at least 5 minutes. As described above, three visual attributes were evaluated in the experiments: brightness, hue and amount of white.

Brightness

A fixed brightness value of 50 was attributed to the achromatic reference stimulus. Just after switching from the test stimulus to the achromatic stimulus, the observers were asked to rate the brightness of the just presented test stimulus relative to the reference achromatic stimulus. Preliminary experiments had shown that it is easier to rate the brightness immediately after the stimulus has disappeared. Furthermore, by showing the reference after each stimulus presentation, any errors due to memory effects were minimized. The following instructions were given, although in Dutch, to each observer:

> *You will see* x [dependent on experiment] *test stimuli. First a reference stimulus will be shown for 5 seconds. Each test stimulus will then be presented for 15 seconds. Between each of these* x *test stimuli, the reference stimulus will again be shown for 5 seconds. You're asked to give a value to the brightness of the test stimulus with respect to that of the reference immediately after the test stimulus disappears. The reference is assigned a brightness value of 50. A value of zero represents a dark stimulus without any brightness. There is no upper limit to the value of brightness, a value of 100 represents a stimulus appearing twice as bright as the reference, a value of 25 is given to a stimulus appearing half as bright, etc.*

Hue and amount of white

When scaling the *amount of white*, observers were asked to assign a percentage of neutral versus coloured to each stimulus. For hue, observers were required to identify the unique hues they could recognize in the coloured part of the stimulus, red – green – yellow – blue, as well as their relative proportions: e.g. 60% red and 40% yellow for a particular orange stimulus. For the hue and amount of white, the following observer instructions were given in Dutch:

> *You will see* x *test stimuli. First a reference stimulus will be shown for 5 seconds. Each test stimulus is then presented for 15 seconds. Between each of these* x *test stimuli, the reference stimulus will be shown again for 5 seconds. Within the 15 seconds the test stimulus is visible, you have to give an answer to the following questions:*
>
> *How much white compared to non-white (or colour) do you recognize in the stimulus? Give a percentage of the amount of white. Keep in mind that this amount of white represents the degree of neutrality. Grey or neutral stimuli can be considered as white. Give 0% when there is only colour visible in the stimulus, give 100% when there is no colour present and thus only a white, a grey or a neutral stimulus is visible.*
>
> *Do you see blue in the stimulus?*
>
> *Do you see green in the stimulus?*
>
> *Do you see red in the stimulus?*
>
> *Do you see yellow in the stimulus?*
>
> *When you see more than one hue, give a percentage to the proportion of each hue present in the stimulus: e.g. 60% red and 40% yellow for a particular orange stimulus.*

When scaling hue, observers are sometimes forced to indicate a maximum of two hues whereby combinations blue-yellow and red-green are not allowed [9]. However, from preliminary experiments, it turned out that some observers do see blue-yellow, red-green or more than two hues in a colour. In order not to force these observers to report things they do not see, observers were free to choose any number of hues and combination.

3.4 Observer variability

The agreement between any two sets of data can be analysed using the coefficient of variation (CV), Eq.(3.1) [50]:

$$CV = 100\sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{(A_i - fB_i)^2}{\bar{A}^2}}$$
$$\text{with } f = \frac{\sum_{i=1}^{n} A_i B_i}{\sum_{i=1}^{n} B_i^2} \tag{3.1}$$

where n indicates the number of data points, A the first dataset and B the second dataset.

For a perfect agreement between two sets of data, this CV value should be zero. The inter-observer agreement for each colour attribute was assessed by calculating the CV values between each individual observer's results and the average of all observers. For the intra-observer agreement the CV values between each individual observer's results of the control stimuli, presented twice to each observer in a single session, were calculated.

Chapter 4

BRIGHTNESS AND THE H-K EFFECT

A couple of days before my first conference, I showed my presentation to one of my supervisors, Peter (P):

I: On this slide, the first results are plotted: a green stimulus with a luminance of 60 cd/m^2 is perceived equally as bright as a white stimulus of 218 cd/m^2.

P: Hmm, I still cannot believe this. Did you double check the measurements?

I: Yes, I did.

P: Can you please check them again with another spectrometer?

…

I: The measurements are validated: green is 60 cd/m^2, white 218 cd/m^2.

P: Fantastic result… but maybe we should do the experiment again with some more observers. Just to be sure…

Expect the unexpected. This chapter deals with the perception of brightness, and its prediction, of stimuli having the same luminance. Although it is expected that they are perceived as approximately equally bright, experiments have shown they aren't…

4.1 Introduction

Luminance is considered as the photometric quantity most closely related to brightness. However, coloured stimuli of equal luminance do not necessarily appear equally bright. The complex relationship between luminance and brightness – conceptualized in a brightness-to-luminance ratio (B/L) – has already been extensively studied [51-54]. Deviations from unity of the B/L ratio have been observed in heterochromatic brightness matches of coloured stimuli. They can be caused by using the erroneous $V(\lambda)$-function in the present standard system of photometry or by a failure of Abney's proportionality and additivity laws [55]. Experimental evidence has shown the latter to be the case in direct heterochromatic brightness matching [56-60]. Although, the relationship between luminance and brightness can be described to a first order approximation by a power law [61], it has become clear that several other parameters such as the luminance of the background and the colourfulness of the stimulus itself are involved as well. The effect of colourfulness or saturation on perceived brightness is, as mentioned before, referred to as the Helmholtz-Kohlrausch (H-K) effect [55], stating that highly saturated colours appear brighter than those of low saturation, even when they are equal in luminance [52, 62].

In the past, a number of models have been developed to describe the brightness of a stimulus. Only six models include the H-K effect: three models based on the concept of equivalent luminance, $L_{\mathrm{Eq,CIE}}$ [53], $L_{\mathrm{Eq,Nay}}$(VCC) and $L_{\mathrm{Eq,Nay}}$(VAC) [52], and three colour appearance models, CAM97u [15], ATD01 [63] and CAMFu [17]. With exception of CAM97u and CAMFu (discussed in Chapter 2), these models are described in the next section. Their predictive performance has been investigated in a couple of psychophysical experiments based on the magnitude estimation method. Due to a severe underestimation of the H-K effect, none of the models performed acceptably. Increasing the weight of the colourfulness contribution to the brightness attribute in the CAM97u model results in a very good correlation between the model predictions and the visually perceived brightness. Finally the experimental results and the brightness prediction obtained from the modified model, referred to as CAM97um, were verified in a matching experiment and a magnitude estimation experiment.

4.2 Vision models for predicting brightness perception

The CAM97u and CAMFu models are fully discussed in Chapter 2. The other four vision models are described below.

ATD01

The ATD01 model is a colour vision model, developed by Guth [63], built on the theoretical ideas of Helmholtz, Hering, von Kries and Mueller [64]. It is developed to predict the brightness, saturation and hue of unrelated colours and to predict a wide range of vision science data on phenomena such as the Bezold-Brücke hue shift, heterochromatic brightness matching, light adaptation and chromatic adaptation [65]. In the ATD model, the *XYZ* tristimulus values are transformed into an *LMS* cone responses. These *LMS* responses are gain-controlled and undergo a second transformation to yield an achromatic (A) and two chromatic or opponent signals (red-green or T, blue-green or D). These A, T and D signals go through a compressive nonlinearity and are finally used to calculate the perceptual attributes brightness, hue and saturation. The brightness Q_{ATD} is calculated as quadrature sum of the A, T and D signals:

$$Q_{\mathrm{ATD}} = \left(A^2 + T^2 + D^2\right)^{0.5} \tag{4.1}$$

As the achromatic A value approximately corresponds to the luminance, this model deals with the H-K effect by adding the chromatic T and D values to the brightness prediction [57].

Equivalent luminance Nayatani ($L_{\mathrm{Eq,Nay}}$)

Brightness has also been modelled based on the concept of equivalent luminance, which is defined as the photopic luminance of a previously determined common reference stimulus that matches the test stimulus (object or self-luminous) in terms of brightness [53]. Three such models have been developed, two by Nayatani ($L_{\mathrm{Eq,Nay}}$) and one by the CIE ($L_{\mathrm{Eq,CIE}}$). The effects of the surround, background and field of view are ignored; the H-K effect is however taken into account. Both models apply in principle only to related colours.

Nayatani proposed two methods that take the H-K effect into account when calculating the equivalent luminance [52]: the Variable-Achromatic-Colour (VAC) and the Variable-Chromatic-Colour (VCC) method, given by Eq.4.2 and Eq.4.3 respectively:

$$L_{\mathrm{Eq,Nay}}(\mathrm{VAC}) = 0.4462L\left[1+\left\{-0.1340q(\theta)+0.0872K_{\mathrm{Br}}(L_{\mathrm{a}})\right\}s_{\mathrm{uv}}(x,y)+0.3086\right]^3 \tag{4.2}$$

$$L_{\mathrm{Eq,Nay}}(\mathrm{VCC}) = 0.4462L\left[1+\left\{-0.8660q(\theta)+0.0872K_{\mathrm{Br}}(L_{\mathrm{a}})\right\}s_{\mathrm{uv}}(x,y)+0.3086\right]^3 \tag{4.3}$$

The presence of the saturation $s_{\mathrm{uv}}(x,y)$ of the test stimulus indicates the inclusion of the H-K effect. The function $q(\theta)$ describes the impact of the hue angle θ on the H-K effect and $K_{\mathrm{Br}}(L_{\mathrm{a}})$ accounts for the increase of the H-K effect when the adapting luminance of chromatic object colours is raised.

In the VAC method, the luminance of the reference achromatic colour is changed in order to match the coloured stimuli. In the VCC method, the luminance of the coloured stimuli is changed in order to match the achromatic reference.

Equivalent luminance CIE ($L_{\mathrm{Eq,CIE}}$)

The CIE equivalent luminance calculates a brightness-related equivalent luminance by using four parameters: the photopic luminance L, the scotopic luminance L', an achromatic adaptation coefficient a, and the chromatic contribution c. The achromatic adaptation contribution a takes the so called Purkinje effect into account. The latter causes a shift in the sensitivity of the human eye towards the blue end of the visible spectrum at low luminance levels. The chromatic contribution c allows specifically for the H-K effect. This chromatic contribution changes with the luminance level and the chromaticity coordinates of the stimulus. A formula for the general equivalent luminance $L_{\mathrm{Eq,CIE}}$ for related colours has been proposed [53]:

$$L_{\mathrm{Eq,CIE}} = (L)^a (L')^{1-a} 10^c \tag{4.4}$$

This equation has been based on visual data gathered from several studies [57, 62, 66-68] and has been tested by matching experiments [54]. It was originally developed based on the 2° quantities, except for the scotopic luminance, but can also be used for a centrally fixed 10° field.

4.3 Experiment 'Lum51'

4.3.1 Psychophysical experiment

The 'Lum51' experiment, named after the luminance value of the coloured stimuli being approximately 51 cd/m^2, was carried out in the darkened viewing room with the LED module as stimulus [23]. For the experiment, 58 coloured stimuli with a constant luminance of 50.97 cd/m^2 (standard deviation 0.80 cd/m^2) and a wide chromaticity gamut were carefully selected (see Fig.4.1). In the experiment, observers were asked to evaluate the perceived brightness, hue and 'amount of white' of the stimuli using the magnitude estimation method, as described above. In this chapter, only the brightness is discussed. The results of the hue and amount of white part are discussed in the next chapter. The experiment started by showing a reference achromatic stimulus with chromaticity close to that of illuminant D65 (u'_{10}, v'_{10}= 0.1979, 0.4695) and a luminance approximately equal to that of the coloured test stimuli, 51.37 cd/m^2. The colour difference $\Delta E_{u'v'}$ between the reference stimulus and the CIE illuminant D65 was 0.0023. To this reference, a fixed brightness value of 50 was attributed. Nine observers, 5 female and 4 male, with ages ranging between 23 and 30 years (average 27) participated in this first experiment. All had normal colour vision according to the Ishihara 24 plate Test for Colour Blindness and were naïve with respect to the purpose of the experiment.

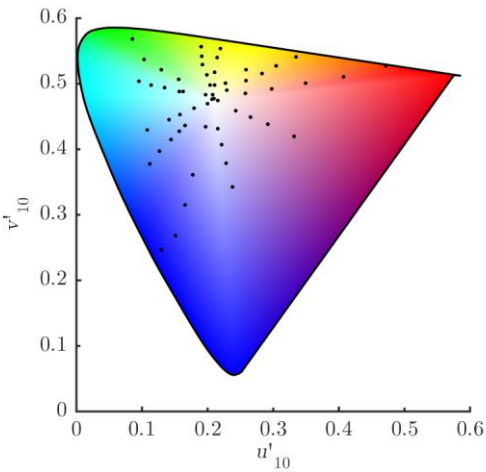


Fig.4.1. Chromaticity coordinates of the 58 stimuli of the 'Lum51' experiment plotted in the *CIE 1976* u'_{10}, v'_{10} chromaticity diagram.

4.3.2 Results for brightness

Observer variability

The values for the inter-observer agreement, assessed as the CV, for brightness in this 'Lum51' experiment ranged from 7% to 21% with an average of 11% and a median of 8%. This result is better than the value of 29% reported by Fu [17] and 40% reported by Koo and Kwak [69] when scaling the brightness of unrelated colours in conditions similar to those used in this study.

Brightness perception

In Fig.4.2, the values of the geometric mean of the observers brightness Q_{avg} for each of the 58 stimuli of equal luminance (51 cd/m²) are plotted against the *CIE 1976* u'_{10}, v'_{10} *saturation* ($s_{uv,10}$), calculated using Eq.4.5 [70]:

$$s_{uv,10} = 13\left[\left(u'_{10} - u'_{n,10}\right)^2 + \left(v'_{10} - v'_{n,10}\right)^2\right]^{1/2} \tag{4.5}$$

Where u'_{10}, v'_{10} and $u'_{n,10}$, $v'_{n,10}$ are the *CIE 1976 chromaticity coordinates* for the CIE 10° observer of the coloured stimulus and the reference achromatic stimulus respectively. Saturation values range from 0.13 to 3.61 (red stimulus).

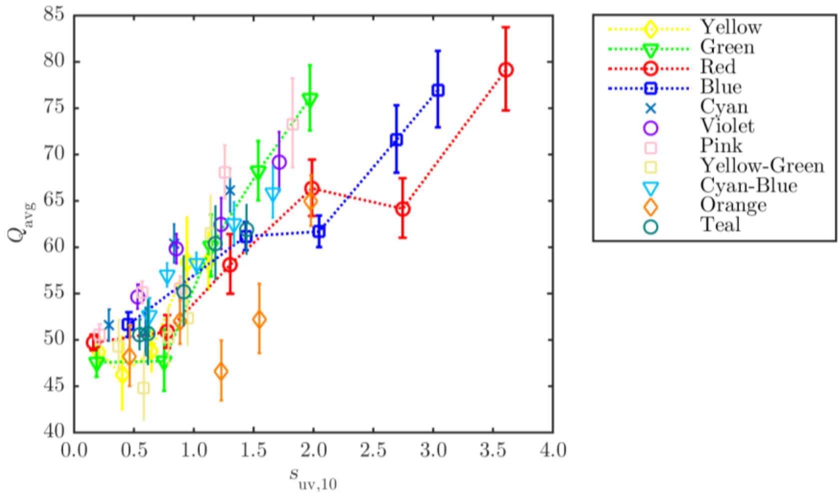


Fig.4.2. Average observed brightness (Q_{avg}) with standard error bars, calculated for each individual stimulus of the 'Lum51' experiment from all observer answers, plotted against the *CIE 1976* u'_{10}, v'_{10} *saturation* ($s_{uv,10}$) of the stimuli.

It is clear that for each hue series, there is generally an increase in perceived brightness with saturation. As each stimulus has the same luminance, Fig.4.2 clearly illustrates the H-K effect.

Model performances

The ability to predict the observed brightness has been investigated for each of the models previously described. Because of the 10° FOV of the test stimuli, the 10° photometric quantities have been calculated, although not all models were developed for this FOV. For each of the six models, the averaged visual brightness as assessed by the observers, Q_{avg}, has been plotted as a function of the predicted brightness on Fig.4.3. The blue, green, red and yellow stimuli have been highlighted.

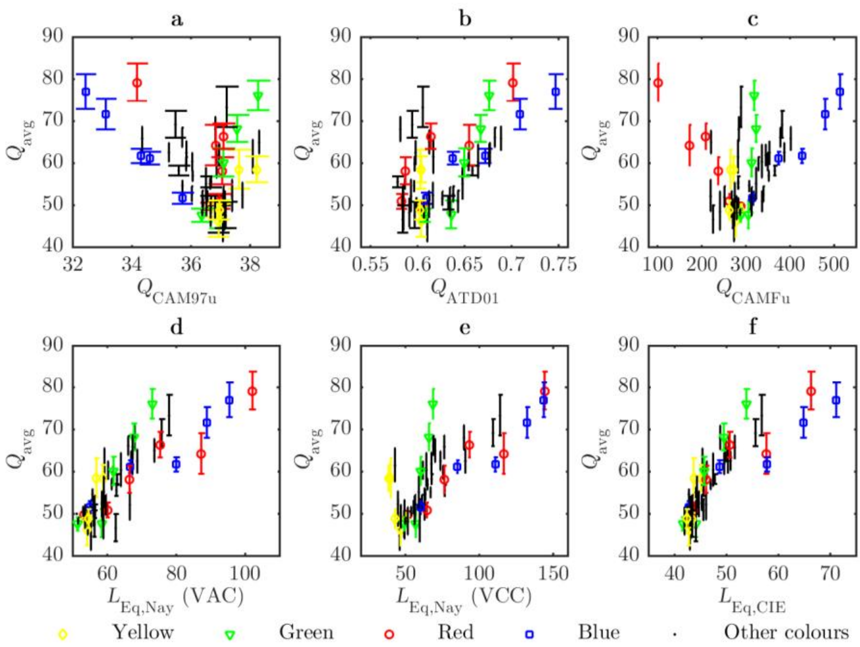


Fig.4.3. Average observer brightness (Q_{avg}) with standard error bars against the brightness predictions of CAM97u (a), ATD01 (b) and CAMFu (c) and the predictions of the equivalent luminance of Nayatani (VAC (d) and VCC (e)) and CIE (f) for the 58 stimuli of the 'Lum51' experiment. The blue, green, red and yellow stimuli are highlighted.

To assess the amount of variation in brightness perception explained by each model, the coefficient of determination (R^2) of the regression between the observed and predicted brightness has been calculated. An R^2 close to 1 suggests a good prediction by the model [71]. Although a linear relation between the observed and the predicted brightness is expected, the Spearman correlation coefficient [71] has also been calculated. The Spearman correlation coefficient r_S is a rank order metric (not sensitive to the potential non-linearity of the relation between observed and predicted values of Q), having a value between -1 (perfect negative correlation) and +1 (perfect positive correlation). Table 4.1 summarizes the values of the R^2 and r_S goodness-of-fit measures for each model.

Table 4.1. Overview of the R^2 and r_S goodness-of-fit measures between the average brightness data of the observers and the predictions of the vision models of the 'Lum51' experiment.

Model	R^2	Spearman r_S
Q_{CAM97u}	0.13	-0.17
Q_{ATD01}	0.34	0.46
Q_{CAMFu}	0.06	0.36
$L_{Eq,Nay}$ (VAC)	0.72	0.88
$L_{Eq,Nay}$ (VCC)	0.60	0.77
$L_{Eq,CIE}$	0.73	0.91

From Table 4.1 and Fig.4.3, it is clear that none of described models perform excellent. Remarkably the two best models, $L_{Eq,CIE}$ (R^2 = 0.73) and $L_{Eq,Nay}$(VAC) (R^2 = 0.72) included the H-K effect explicitly, but were not developed for unrelated colours. Furthermore, the VAC model of Nayatani's equivalent luminance, where the achromatic stimulus is changed to match the coloured stimuli, performs better than the VCC model (R^2 = 0.60) although it should be less applicable to the method used in this experiment, using a constant achromatic stimulus.

Although CAM97u and CAMFu have been developed specifically for unrelated colours and include explicitly the H-K effect, the low values of the Spearman correlation coefficient and the low coefficient of determination (Table 4.1) indicate that they are unable to predict the experimental brightness data (with a luminance of 51 cd/m^2). The brightness prediction of CAM97u (Eq.2.19) and CAMFu (Eq.2.23) both consist of a summation of an achromatic signal, which is nearly constant (all stimuli have equal luminance), and a contribution of the colourfulness factor M, which takes into account the H-K effect. The failure of these two models to predict the perceived brightness for the self-luminous stimuli might be attributed to a poor implementation of the colourfulness factor. This is suggested by the good correlation between M_{CAM97u} and Q_{avg} (R^2 = 0.77 and r_S = 0.86). This correlation indicates that the predictive performance of the brightness of

CAM97u (Chapter 2, Eq.2.19) could be considerably improved just by increasing the contribution of the colourfulness:

$$Q_{\mathrm{CAM97u}} = \left[(1.1)\left(A_{\mathrm{CAM97u}} + w_{\mathrm{M}} \times M_{\mathrm{CAM97u}}\right)\right]^{0.9} \tag{4.6}$$

Where Q_{CAM97u} is the predicted brightness, A_{CAM97u} the achromatic signal, M_{CAM97u} the colourfulness and w_{M} a weighting factor - equal to 0.01 in CAM97u - that regulates the contribution of colourfulness to the brightness prediction [16].

Because A_{CAM97u} and M_{CAM97u} are not totally independent quantities, additional visual data at different luminance levels were required in order to determine the weighting factor and to propose an improved model.

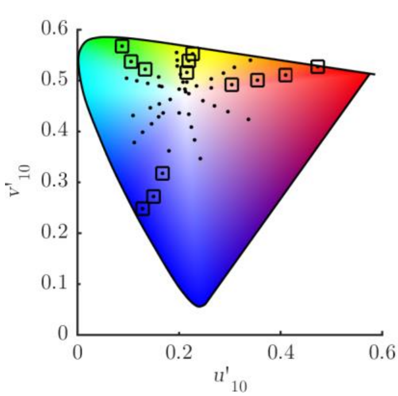
Note that, while the colourfulness factor M_{CAM97u}, calculated according to CAM97u, is relatively well correlated to the brightness, the correlation between the colourfulness calculated according to CAMFu and Q_{avg} (R^2 = 0.48 and r_{S} = 0.61), indicates a rather poor prediction by this model. The failure of the CAMFu model might be attributed to the use of a colourfulness factor M_{CAMFu} based on the CIECAM02 model [12], which was never intended for unrelated colours.

4.4 Experiment 'Lum6'

4.4.1 Psychophysical experiment

In 'Lum6', a second magnitude estimation experiment performed in the darkened viewing room with the LED module as stimulus and named after the luminance value of the coloured stimuli being approximately 6 cd/m^2, observers were asked to evaluate the brightness, hue and amount of white of unrelated self-luminous stimuli [72]. Again, in this chapter only the evaluation of brightness is discussed, the results of the hue and amount of white experiments are discussed in the next chapter.

To improve the brightness prediction of CAM97u, particularly to take the H-K effect correctly into account, coloured stimuli each having more or less the same value of A_{CAM97u} are preferred. Therefore a set of 58 coloured stimuli with a FOV of 10° and equal luminance have been selected. However, this time a photopic luminance level of 6.23 cd/m^2 (standard deviation 0.11 cd/m^2) was chosen to extend the validity of the results found in the first experiment 'Lum51'. The CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the stimuli are illustrated in Fig.4.4 (left). In addition to these 58 coloured stimuli, a set of 17 achromatic stimuli with a luminance between 7.54 cd/m^2 and 47.60 cd/m^2 and a chromaticity close to that of illuminant D65 (u'_{10}, v'_{10} = 0.1979, 0.4695; mean $\Delta E_{u'v'}$=0.0053) has been selected (see Fig.4.4, right). The colourfulness of these achromatic stimuli is very low and approximately the same. These stimuli enable to obtain a single brightness scale appropriate to both chromatic and achromatic stimuli. The luminance of the reference stimulus was approximately the same as the luminance of the coloured stimuli, 6.38 cd/m^2, and the chromaticity was close to that of illuminant D65 ($\Delta E_{u'v'}$ = 0.0063).



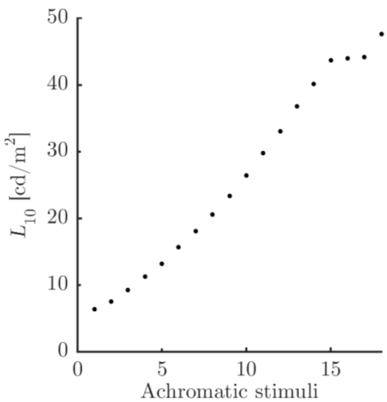


Fig.4.4. (left) CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the 58 coloured stimuli of the 'Lum6' experiment. The stimuli highlighted with a squared symbol are also used in a validation experiment 'Match' described below. (right) Luminance of the 17 achromatic stimuli of the 'Lum6' experiment, calculated using the CIE 10° observer.

Twenty observers, 10 male and 10 female, with ages ranging between 20 and 31 years (average 25) participated in the experiment. All had normal colour vision according to the Ishihara 24 plate Test for Colour Blindness. Six observers already participated in the previous experiment 'Lum51' while the others were naïve with respect to the purpose of the experiment.

4.4.2 Results for brightness

Observer variability

In this experiment, in addition to the inter-observer agreement both short term and long term intra-observer agreement was also assessed. The short term intra-observer agreement was analysed by randomly selecting ten stimuli and having each observer rate them a second time at the end of the experiment. This short term agreement was quantified by calculating the CV values between each individual observer's results of the ten stimuli during the test and their results of the same ten stimuli at the end of the test. The observers were not aware that ten stimuli were presented a second time. The long term intra-observer agreement was analysed by having three observers (2 male, 1 female) repeat the experiment three months later and was quantified by calculating the CV values between each individual observer's results of both experiments. Observers were not told that the experiment was identical.

The results for the inter- and intra-observer agreement are summarized in Table 4.2 in terms of CV values. These results show that the mean CV values for inter-observer, short term intra-observer and long term intra-observer agreement are 13%, 11% and 8%, respectively. The CV values are fairly low for all observers, indicating a good agreement. The values for inter-observer agreement are much better than the value of 29% reported by Fu et al. [17] and the 40% reported by Koo and Kwak [69] and similar to the value of 11% reported above in the 'Lum51' experiment. The CV value for the short term intra-observer agreement of 11% could only be compared to the 15% 'repeatability' obtained by Fu et al. [17], as none of the other studies reported an intra-observer agreement.

Table 4.2 The inter-observer, short term intra-observer and long term intra-observer agreement of the 'Lum6' experiment as assessed by the coefficient of variation CV (%).

Observer #		Inter-observer agreement		Short term intra-observer agreement		Long term intra-observer agreement	
1	2	12	10	8	9	6	-
3	4	12	11	13	6	-	-
5	6	15	11	19	2	-	-
7	8	13	10	10	13	-	-
9	10	9	12	10	17	-	-
11	12	13	23	13	10	-	-
13	14	9	8	12	7	-	-
15	16	10	12	11	11	7	12
17	18	25	15	11	12	-	-
19	20	23	18	21	10	-	-
Mean		**13**		**11**		**8**	
Median		**12**		**11**		**7**	

Brightness perception

Again, it is clear that for each hue series, there is generally an increase in perceived brightness with saturation (see Fig.4.5). As each stimulus has the same luminance, this figure clearly illustrates the H-K effect. In fact the slopes of most hue series seem to be coincident, suggesting that the effect of saturation on brightness is equal for most hue series. This can be checked using a customized ANCOVA. However, before calculating this ANCOVA, it should be shown that the observed brightness is indeed significantly different between these equi-luminance coloured stimuli.

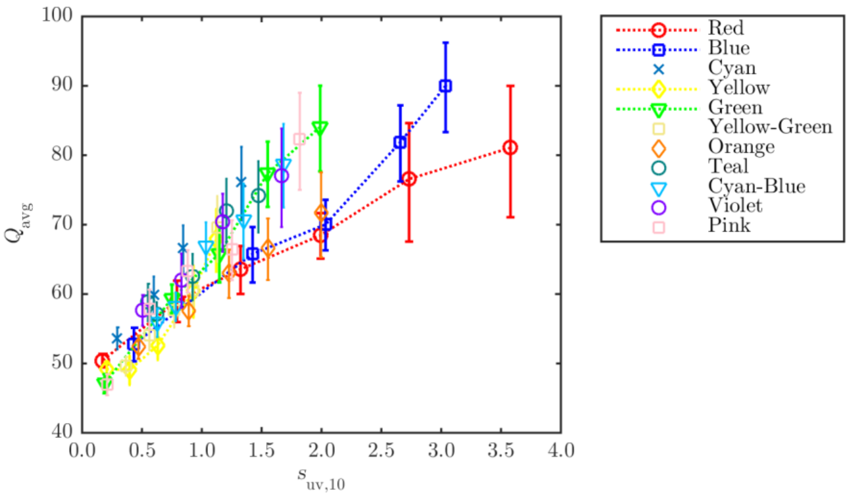


Fig.4.5. 'Average observer' brightness (Q_{avg}) with standard error bars, calculated for each individual coloured stimulus of the 'Lum6' experiment from all observer answers, plotted against the CIE 1976 u'_{10}, v'_{10} *saturation* ($s_{uv,10}$).

As the same group of observers rated the brightness of each stimulus, a one-way repeated-measures design of ANOVA on all coloured stimuli was calculated. The analysis showed that, although all coloured stimuli have the same luminance, their brightness perception was significantly different between each other, $F(1.824, 34.659) = 14.801$, $p < 0.001$. A customized ANCOVA with Q_{avg} as dependent variable, the 11 hues as fixed factors and $s_{uv,10}$ as covariate, showed that the effect of the interaction term between Q_{avg} and $s_{uv,10}$ is significant, $F(10,1058) = 2.155$, $p < 0.05$, while the same analysis using only 9 hues (without red and blue) as fixed factors is not significant, $F(8,842) = 0.485$, $p = 0.867$. This indicates that the regression slopes are homogeneous for all colours except for red and blue. Although some studies reported that the H-K effect is different or even absent for yellow [55], this extensive study suggests that the H-K effect, which is clearly visible, is only different for red and blue.

Remarkably, four of the twenty observers, although obtaining good results in the Farnsworth Munsell 100 Hue Test [47], rated the red stimuli as being less bright compared to the reference stimulus. Although the other colours were rated brighter and the CV values of these observers were normal, these four observers indicated, independently from each other, having trouble with rating the brightness of red stimuli. When converting the brightness values to *z*-scores using SPSS [71, 73], it seemed that none of these results are outliers, so these results were not removed from the experiment.

Model performances

The brightness predictions according to the six vision models described earlier were again compared to the 'average observed' brightness of the stimuli. In Fig.4.6, the perceived brightness, Q_{avg}, has been plotted against the predicted brightness for each of these models. This figure clearly indicates a different slope for the coloured (black x's and coloured symbols) and the achromatic stimuli (black circles). In table 4.3 the statistical results for each model obtained with all stimuli and with the achromatic stimuli only, are summarized.

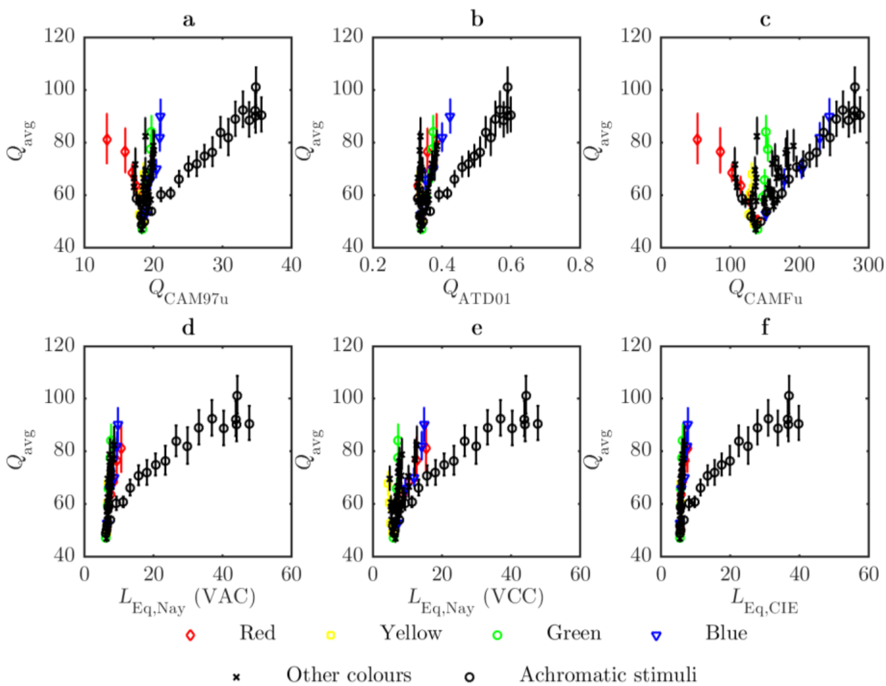


Fig.4.6. Perceived brightness (Q_{avg}) with standard error bars against the brightness predictions of CAM97u (a), ATD01 (b) and CAMFu (c) and the predictions of the equivalent luminance of Nayatani (VAC (d) and VCC (e)) and CIE (f) for the 75 stimuli of the 'Lum6' experiment.

In Table 4.3, the low values of the Spearman correlation coefficient r_S and the low coefficient of determination R^2 for all stimuli are striking. It is clear that none of the described models performs satisfactorily, in accordance with the conclusion of the previous 'Lum51' experiment. However, as is also visible in Fig.4.6, when only the achromatic stimuli are considered (see Table 4.3, achromatic stimuli), it can be observed that all models perform well. This indicates that the low overall correlation is due to a severe underestimation of the H-K effect.

Table 4.3. Overview of the R^2 and r_S goodness-of-fit measures between the 'average observed' brightness data and the predictions of the vision models of the 'Lum6' experiment.

Model	All stimuli		Achromatic stimuli	
	R^2	Spearman r_S	R^2	Spearman r_S
Q_{CAM97u}	0.47	0.61	0.95	0.95
Q_{ATD01}	0.57	0.63	0.95	0.95
Q_{CAMFu}	0.47	0.57	0.95	0.95
$L_{Eq,Nav}$ (VAC)	0.53	0.87	0.91	0.95
$L_{Eq,Nav}$ (VCC)	0.59	0.78	0.91	0.95
$L_{Eq,CIE}$	0.51	0.86	0.91	0.95

Modified CAM97u: CAM97um

As suggested earlier, the brightness prediction of CAM97u, Eq.4.6, could be improved by increasing the colourfulness weighting factor w_M, taking the H-K effect better into account. To determine an optimized colourfulness weighting factor, the perceived brightness Q_{avg} was first rescaled to the original CAM97u ($w_M = 0.01$) brightness scale but using only the data of the achromatic stimuli for which the CAM97u model seems to be working well. This 'rescaled observed' brightness, Q_r, was obtained by multiplying Q_{avg} with the slope of the linear regression between Q_{avg} and Q_{CAM97u} for all 17 achromatic stimuli. By minimizing the mean of the squared residual errors between Q_r and the brightness values calculated according to Q_{CAM97u} (Eq.4.6), the value of the colourfulness weighting factor w_M was optimized from its original value of 0.01 to 0.268. Similar to Eq.4.6, the modified brightness model, $Q_{CAM97um}$, is then given by:

$$Q_{CAM97um} = \left[(1.1)\left(A_{CAM97u} + 0.268 \times M_{CAM97u}\right)\right]^{0.9} \tag{4.7}$$

When plotting the 'average observed' brightness Q_{avg} against $Q_{CAM97um}$ for all stimuli, it is clear that the new model outperforms the former models (see Fig.4.7). This is confirmed by the high Spearman correlation coefficient ($r_S = 0.96$) and the high coefficient of determination ($R^2 = 0.91$).

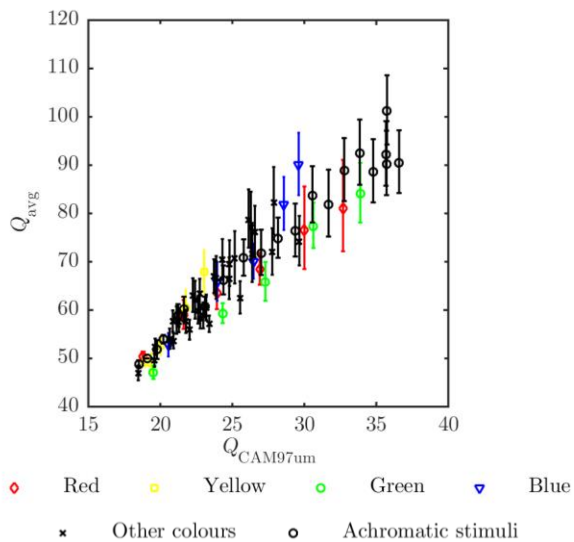


Fig.4.7. Perceived brightness (Q_{avg}) with standard error bars plotted against the brightness predictions of CAM97um for the 75 stimuli of the 'Lum6' experiment.

As mentioned before, a colour appearance model can be considered successful when the error of the model's prediction is smaller than the observer accuracy in terms of inter-observer agreement [17]. The goodness-of-fit of the CAM97um model for predicting the brightness, as assessed by the coefficient of variation (CV = 6%), is substantially lower than the inter-observer variability (CV = 13%) and lower than the goodness-of-fit of the other six vision models (CV between 14 and 59), which proves again its excellent performance.

4.5 Validation

4.5.1 Experiment 'Match'

The performance of the modified CAM97um model is verified by a successive matching experiment, called 'Match', performed by the same observers as in the 'Lum6' experiment. The matching experiment started immediately after the 'Lum6' experiment except for a break of 15 minutes. In the experiment, which lasted for about 25 minutes, observers adjusted the intensity of the achromatic reference stimulus until it matched that of the coloured stimulus in terms of brightness. In this experiment, no evaluation of hue and amount of white was performed. From the 58 coloured stimuli, only the four most saturated red, and the three most saturated blue, yellow and green stimuli have been used (see Fig.4.4 (left)). The initial luminance of the reference stimulus, shown in temporal juxtaposition with the coloured stimuli, was randomly high or low in order to avoid an initial luminance bias [74, 75]. Observers were able to switch back and forth between the reference and the coloured stimulus as much as they wanted to until a satisfactorily match was found. As in the magnitude estimation experiment, the coloured stimulus was always shown for 15 seconds. Two groups of 10 observers viewed the same sequence of coloured stimuli but with an opposite initial reference luminance.

The 10° luminance of the reference was measured after each match and an "average matched reference luminance" was obtained for each coloured stimulus by taking the arithmetic mean of all observer matches. A high initial luminance of the reference mostly resulted in a higher matched reference luminance compared with a low initial luminance (see Table 4.4). This effect is responsible for an average luminance difference of 22% between the two experimental conditions. However, the experiment was set up with both conditions having an equal number of matches. By averaging the results, this type of bias was neutralized [44].

Table 4.4. Initial luminance bias: values of the matched reference luminance of the 13 coloured stimuli for both the high and low initial luminance and the difference between them, ordered by hue.

Colour	Matched luminance (cd/m^2)		Difference (cd/m^2)
	High initial luminance (H)	Low initial luminance (L)	H-L
Red	32.3	29.1	3.1
	23.2	20.3	2.9
	15.9	16.1	-0.3
	15.1	10.4	4.7
Blue	32.1	27.1	5.0
	22.2	25.8	-3.6
	23.2	15.7	7.4
Yellow	15.5	7.9	7.6
	10.7	9.7	1.0
	10.4	6.4	3.9
Green	23.3	18.3	5.1
	20.8	15.7	5.0
	20.1	10.2	9.9
Mean	**18.4**		**4.0**

A plot of the averaged matched reference luminance versus the saturation $s_{uv,10}$ of the stimuli that were to be matched, is given in Fig.4.8 (left). Although the luminance of the coloured stimuli were all approximately 6 cd/m^2, the figure indicates that the most saturated red and blue ones were matched for their brightness with an achromatic stimulus having a luminance of around 30 cd/m², clearly illustrating the H-K effect. In Fig.4.8 (right) a plot of the modified CAM97um brightness prediction of the matched reference against the prediction of the corresponding stimuli is given. The figure indicates that the modified CAM97um is capable to predict the brightness outcome of the matching experiment. The low value of 7% for the CV value between both brightness predictions confirms the excellent performance of CAM97um.

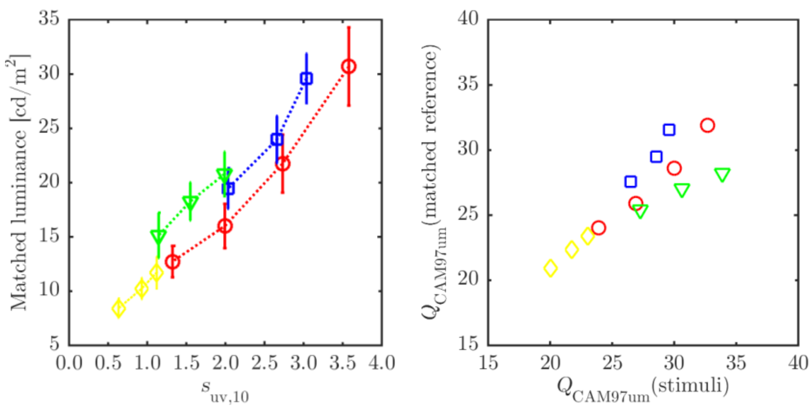


Fig.4.8. (left) Matched reference luminance with standard error bars plotted against the CIE 1976 u'_{10}, v'_{10} *saturation* ($s_{uv,10}$) of the corresponding stimuli. (right) The modified CAM97um brightness prediction of the matched reference plotted against the prediction of the corresponding stimuli.

4.5.2 Experiment 'Random'

A decisive magnitude estimation experiment to validate the modified CAM97um brightness prediction was set up with 107 stimuli: 15 achromatic stimuli, 40 coloured stimuli and 52 'random' stimuli. In this experiment, no evaluation of hue and amount of white was involved. The luminance of the achromatic stimuli ranged from 5.94 cd/m² to 297.47 cd/m² (see Fig.4.9 (left)) with a chromaticity close to that of illuminant D65 (mean $\Delta E_{u'v'}$ = 0.0022). The 40 coloured stimuli consisted of the four primary hues with both a low and a high saturation, at five luminance levels (see Fig.4.9 (left)). The 52 'random' stimuli had a luminance ranged randomly within 6.48 and 57.60 cd/m² (see Fig.4.9 (middle)) and covered the whole chromaticity gamut of the experimental setup (see Fig.4.9 (right)).

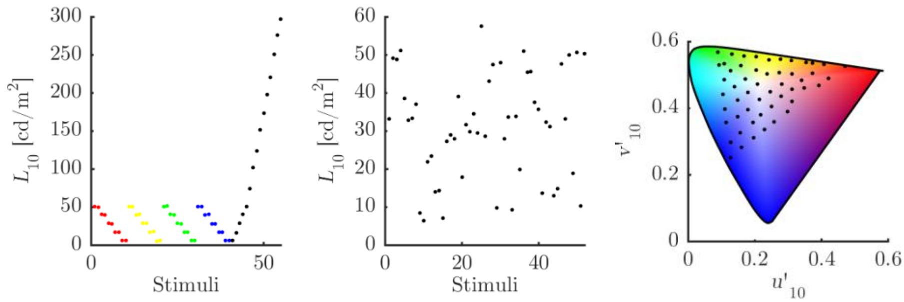


Fig.4.9. (left) Luminance values of the 40 coloured and the 15 achromatic stimuli. (middle) Luminance values of the 52 'random' stimuli. (right) CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the 52 'random' stimuli.

The experimental method used in this 'Random' experiment was identical to the one used in the magnitude estimation experiments described above, except for the reference stimulus to which an intermediate luminance of 43.10 cd/m^2 was attributed. Twenty observers participated in this experiment. All except two had also participated in the 'Lum6' and matching experiment. The mean CV values for inter-observer, short term intra-observer and long term intra-observer agreement of this validation magnitude estimation experiment are 18%, 12% and 15%, respectively, and are similar to the values mentioned before.

The geometric mean was again used to obtain the observer brightness Q_{avg} ("average observer") and was plotted versus Q_{CAM97um} in Fig.4.10. From the coefficient of determination of 0.81, the Spearman correlation coefficient of 0.90, and the CV value of 11%, it is clear that the modified CAM97um model gives an excellent prediction of brightness, given by Eq.4.7, of both coloured and achromatic unrelated self-luminous stimuli covering a wide colour gamut and range of luminance levels.

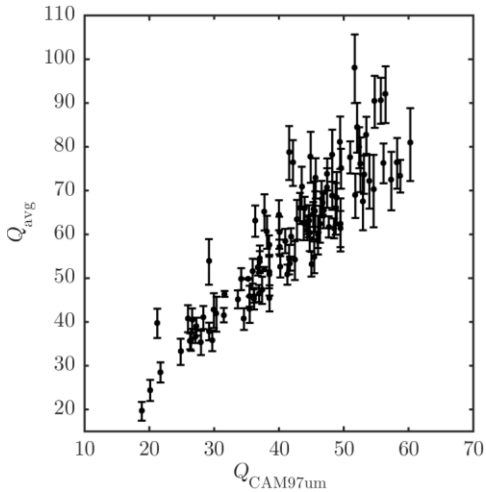


Fig.4.10. Perceived brightness (Q_{avg}) with standard error bars plotted against the brightness prediction of CAM97um.

4.6 Conclusions

In several psychophysical experiments, the brightness perception of unrelated self-luminous coloured stimuli was investigated. First, 58 coloured stimuli having a constant luminance level of 51 cd/m^2 were investigated in a magnitude estimation experiment with 9 observers. Next, coloured stimuli having a constant luminance level of 6 cd/m^2 and a set of achromatic stimuli having a luminance ranging from 8 cd/m^2 to 47 cd/m^2, were investigated in a second magnitude estimation experiment with twenty observers. It was found that the Helmholtz-Kohlrausch effect contributed significantly to the observed brightness. The ability of six vision models to predict the observed brightness was evaluated using the coefficient of determination and the Spearman correlation coefficient. Although the models included the H-K effect and three of them were developed particularly for unrelated colours, none of the models seemed to be able to predict the perceived brightness satisfactorily. The expected linear relationship between the observed and predicted brightness was not achieved.

Adapting the CAM97u model by increasing the colourfulness contribution in the brightness attribute resulted in modified model, called CAM97um, which allows for a substantially better brightness prediction. The performance of the new model was confirmed by both a matching experiment and an extensive validation magnitude estimation experiment using a random sequence of stimuli within a wide chromaticity range, including achromatic ones, and within a wide range of luminance values. The modified model CAM97um clearly outperformed the other existing vision models and was found to give a reliable brightness prediction for unrelated self-luminous stimuli.

Chapter 5

A NEW CAM FOR UNRELATED SELF-LUMINOUS STIMULI: CAM15u

Trying to convince me to work harder, Peter (my supervisor) often asked me about the 'Withouck' model. As months passed by, the model was steadily developed and I finally gave it a name: CAM15u. Job done? Of course not, refinements were needed and after each 'finalised version', discussions between Peter, Kevin and myself resulted in a new idea: "The achromatic part of the brightness scale should be fixed on the old CAM97u brightness scale.", "It's more logical to represent the colourfulness by the strength of the colour of the stimulus.",… Meanwhile, as the model was given a name, Peter lost his trigger. However, it wasn't long before he came into my office and asked me about 'one Martijn', the unit of the brightness scale of CAM15u…

This chapter is about, what I consider to be my biggest scientific accomplishment, a new CAM for unrelated self-luminous stimuli: CAM15u.

5.1 Introduction

In the previous chapter, it has been shown that the CAM97u and CAMFu models were unable to accurately predict the perceived brightness of unrelated self-luminous stimuli [23], mainly due to an underestimation of the H-K effect. In the former chapter, a modified model was proposed, CAM97um, that substantially improved the brightness prediction of CAM97u by simply increasing the weight of the colourfulness contribution to brightness [72].

In this chapter, an entirely new CAM for unrelated self-luminous colours, CAM15u, is presented [76]. The main features of the model are the use of the absolute spectral radiance of the stimulus as input, the use of the CIE 2006 cone fundamentals [53], the inclusion of the Helmholtz-Kohlrausch effect, the *amount of white* as an alternative perceptual attribute to saturation and colourfulness and a simplified calculation procedure compared to existing models. The model predicts the brightness, hue, colourfulness, saturation and the amount of white. The CAM15u model is restricted to photopic, non-glare-inducing unrelated stimuli having a field of view of 10°. The model has been developed and validated using data obtained in a magnitude estimation experiment in which twenty observers have rated more than 150 unrelated self-luminous stimuli for three absolute perceptual attributes: brightness, hue and amount of white. This new CAM for unrelated colours, CAM15u, is shown to be accurate and to outperform existing models.

5.2 Psychophysical Experiment

5.2.1 Experimental setup

The visual experiments described in this chapter were carried out in the darkened viewing room with the LED module as stimulus. To create and validate a new CAM for unrelated self-luminous colours, a test set of 105 stimuli and a validation set of 52 stimuli were carefully selected (see Fig.5.1). These stimuli were chosen to cover a large portion of the chromaticity diagram. Their 10° luminance values were randomly selected from a 6.21 to 56.61 cd/m² luminance range, which provides photopic stimulus viewing conditions while avoiding glare.

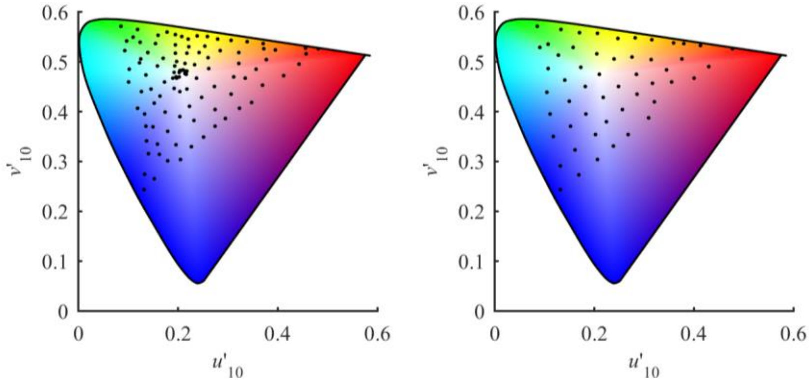


Fig.5.1. CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the 105 test stimuli (left) and 52 validation stimuli (right).

5.2.2 Visual attributes

As mentioned before, the colour appearance of a scene is described by absolute attributes such as brightness, colourfulness, and hue and by relative attributes such as lightness, chroma, and saturation [1]. For self-luminous stimuli, brightness, hue, colourfulness and saturation are most relevant. However, as mentioned, it turned out that instead of rating colourfulness, observers were more comfortable evaluating the "*amount of white versus non-white*" in a stimulus.

5.2.3 Experiment procedure

Twenty observers, 9 female and 11 male, with ages ranging between 21 and 32 years (average 24.5) participated in the psychophysical experiment. All had normal colour vision according to the Ishihara 24 plate Test for Colour Blindness and the Farnsworth-Munsell 100 Hue Test (mean Total Error Score of 31, indicating all observers had an average or superior discrimination) [47]. Thirteen of them had participated in one or more of the previous experiments, while the others were naïve with respect to the purpose of the experiment. Prior to the experiment, observers adapted to the dark viewing conditions.

To reduce the influence of fatigue in the experiment, the combined set (test and validation) of stimuli was presented in two sessions taking about 35 minutes each. In each session, about 90 stimuli were presented: 10 control stimuli to estimate the intra-observer accuracy and about 80 randomly chosen stimuli from the test and validation sets. A break of about 15 minutes was offered between each session. For each session, the stimuli were randomly arranged in two series, each being evaluated by half of the observers to avoid possible bias due to the series sequence [49]. Also, as preliminary experiments have shown that observers have difficulty in rating all three attributes at once, the brightness was rated separately from the hue and amount of white. About half of the observers started with scaling their two sessions for brightness, while the other half started with scaling the hue and amount of white. Each stimulus was presented to the observers for 15 seconds. Between these stimuli, the reference achromatic stimulus was shown for 5 seconds.

When scaling brightness, the stimuli were rated in comparison with a 51.20 cd/m² reference achromatic stimulus shown in temporal juxtaposition and to which a brightness value of 50 was attributed. The 10° luminance of this reference achromatic stimulus was chosen to correspond to a perceived brightness (as calculated by the CAM97um model) approximately midway the brightness range of all the stimuli in this experiment. The chromaticity of the reference stimulus (u'_{10}, v'_{10} = 0.2111, 0.4750) was close to that of the equi-energy stimulus, EEW (u'_{10}, v'_{10} = 0.2105, 0.4737; $\Delta E_{u',v'}$ = 0.0014).

5.3 Observer data

For each attribute, the inter- and intra-observer variability was assessed using the coefficient of variation (CV), Eq.3.1 [50].

5.3.1 Brightness

The low average inter-observer CV values for the test and validation set, respectively 17% and 14%, as well as the small CV range (see Table 5.1), indicate that observers agreed well and had little difficulties in scaling brightness. The intra-observer variability had an average CV value of 20%. The mean CV values for inter-observer variability are comparable to those reported above (11%, 13% and 18%), and better than the ones reported in Fu et al. [17] and in Koo and Kwak [69] (respectively 29% and 40%). All studies had similar conditions. The mean CV value for intra-observer agreement is slightly higher than the 15% repeatability obtained by Fu et al. [17] and the 11% and 12% short-term intra-observer agreement reported above.

As proposed by ASTM International [43], the perceived brightness scaling for an average observer, $Q_{avg,i}$, was again obtained by calculating the geometric mean of all the observers' brightness scaling $Q_{obs,i}$ for each stimulus *i*.

Table 5.1. Inter- and intra-observer agreement for the test and validation set in terms of the coefficient of variation CV (%).

Coefficient of Variation CV (%)	Brightness			Hue			Amount of white		
	Test	Val.	Both	Test	Val.	Both	Test	Val.	Both
Observer #	Inter		Intra	Inter		Intra	Inter		Intra
1	18	14	14	15	15	9	23	29	24
2	21	19	22	10	9	12	21	27	32
3	16	14	26	9	9	10	25	29	46
4	19	11	23				29	42	46
5	20	24	32						
6	17	13	23	9	8	12	49	71	48
7	23	11	13	10	8	9	34	35	49
8	18	12	12	8	7	10	23	21	31
9	10	10	14	9	22	8	38	39	85
10	21	14	6	7	8	10	31	33	40
11	10	8	6	8	8	8	24	27	22
12	11	9	12	10	9	8	20	34	37
13	24	19	33				38	37	47
14	14	12	17	10	7	12	28	34	41
15	14	11	23	12	7	8	26	40	75
16	14	11	10	7	8	12	31	42	40
17	22	21	40	10	9	17	44	45	73
18	20	11	19	17	24	6	37	36	33
19	21	14	25	12	20	18	21	28	38
20	15	15	25						
Mean	**17**	**14**	**20**	**10**	**11**	**11**	**30**	**36**	**44**
Median	18	13	21	10	8	10	28	35	40

5.3.2 Hue (quadrature)

For hue, a quadrature scale was developed by transforming all the observers' results into a 0-400 scale [12, 31]: 0-100 for red-yellow, 100-200 for yellow-green, 200-300 for green-blue and 300-400 for blue-red. For example, a value of 40 for the hue of a particular orange stimulus containing 60% red and 40% yellow. Stimuli with a median amount of white above 90 were excluded from the analysis as most observers had difficulty recognizing hue, let alone their relative proportions, in these stimuli.

As mentioned above, observers were not restricted in the number or combination of perceived unique hues they could report. Although binary combinations of blue-yellow and red-green or combinations of three or four hues cannot be transformed into a 0-400 scale, and were therefore excluded from the experiment, non-forced evaluation does provide interesting information about the actual perception of observers. For 16 of the 20 observers these cases almost never occurred: out of the 2512 answers a red-green combination was reported once and a yellow-blue combination 21 times. However four observers, i.e. 20%, showed very divergent responses using the non-forced evaluation. This suggests that a hue quadrature scale is not as representative of typical hue perception as commonly believed. Observer 5 and 20 indicated having trouble with scaling hue as they were always thinking about mixing paints and found themselves unreliable for scaling the attribute, observer 4 perceived a yellow hue in almost all the stimuli and thus often perceived blue and yellow together, and observer 13 often reported perceiving more than two hues in a presented stimulus. Although these four observers obtained good results in the Farnsworth Munsell 100 Hue Test and were very dedicated to their task, their answers could not be mapped to the hue quadrature scale traditionally used in CAMs. For this reason, they could not be used in this study and thus all their hue related results were excluded from the analysis. The mean inter-observer CV values for the 16 other observers for the test and validation set were respectively 10% and 11%. The average intra-observer CV was 11%. These low CV values, for all observers (see Table 5.1), indicate a good agreement. Several studies with similar experimental conditions reported comparable levels of agreement: 9% by Luo et al. [9], 12% by Koo and Kwak [69] and 15% by Fu et al. [17] for inter-observer agreement and 6% by Fu et al. [17] for intra-observer agreement.

By calculating the arithmetic mean of all the observers' hue quadrature responses $H_{\mathrm{obs,i}}$ for each stimulus i (with four outliers excluded) an average observer perceived hue quadrature $H_{\mathrm{avg,i}}$ was obtained.

5.3.3 Amount of white

The CV values for the amount of white were only calculated for 18 observers (see Table 5.1) as two observers indicated having trouble with scaling the amount of white and their answers diverged substantially from the bulk of the observer answers. The mean inter-observer CV values (with two outliers excluded) for the test and validation set are respectively 30% and 36%. The mean intra-observer CV was 44%. These inter-observer values are typical for this kind of attribute, e.g. values of 27% and 39% were found by Koo and Kwak [69] and by Fu et al. [17], respectively, for the colourfulness of unrelated colours. Although the amount of white was expected to be a more familiar attribute than the colourfulness, it generally does not lead to a more robust estimate. This is partly the result of the high difficulty in quantifying the amount of white for saturated stimuli. However, because of its familiarity and simplicity, amount of white is still considered as the preferred attribute in this experiment.

As the distribution of the ratings of the amount of white becomes more skewed near the fixed end points (0% and 100%), the median of the observers' amount of white $W_{obs,i}$ for each stimulus i (with two outliers excluded) was calculated to obtain an average observer perceived amount of white $W_{avg,i}$.

5.4 Development of CAM15u

Following the current understanding of human colour perception, based on the results of the psychophysical experiment and inspired by other CAMs, such as CAM97u and CAMFu, a new parametrically simpler and more accurate model to predict the colour appearance of unrelated self-luminous colours, CAM15u, has been developed. In what follows, the various steps of the model, as well as critical differences with previous CAMs, are discussed.

5.4.1 Absolute, normalized cone excitations

As mentioned in Chapter 1, human colour vision starts with light absorption by the photo-sensitive receptor cells - the rods and cones - in the retina. The cones, dominating photopic vision, come in three different types and are typically referred to as the ρ, γ, β cones. In basic colorimetry, the colour of a stimulus is usually specified in terms of the CIE tristimulus values XYZ. The latter are calculated from the CIE colour matching functions (CMF), which are based on matching data obtained in psychophysical experiments using either 2° or 10° stimuli. They can be linearly transformed to LMS type CMFs, called cone fundamentals. The latter are the effective cone excitations taking into account the spectral absorption characteristics of the ocular media and the macular pigment, and the self-screening in the outer segment of the photoreceptors. Recently, the CIE provided a new set of cone fundamentals specifically suited to 10° stimuli [77]. These cone fundamentals were derived from the best set of colour-matching functions experimentally collected on a 10° field [32, 78, 79]. Although the use of the CIE 2006 cone fundamentals do not significantly change the results compared to the use of CIE 1964 XYZ CMF's (see Section 5.9), they are the most recent fundamentals as proposed by the CIE. In the CAM15u model they are used to calculate the fundamental cone excitations, ρ_{10}, γ_{10}, β_{10}, of the stimulus:

$$\rho_{10} = k_\rho \int_{390}^{830} L_{e,\lambda}(\lambda)\bar{l}_{10}(\lambda)\,d\lambda$$
$$\gamma_{10} = k_\gamma \int_{390}^{830} L_{e,\lambda}(\lambda)\,\bar{m}_{10}(\lambda)\,d\lambda \tag{5.1}$$
$$\beta_{10} = k_\beta \int_{390}^{830} L_{e,\lambda}(\lambda)\,\bar{s}_{10}(\lambda)\,d\lambda$$

With λ the wavelength from 390 to 830 nm, $L_{e,\lambda}(\lambda)$ the spectral radiance of the stimulus and $\bar{l}_{10}(\lambda)$, $\bar{m}_{10}(\lambda)$ and $\bar{s}_{10}(\lambda)$ the CIE 2006 10° cone fundamentals in terms of energy [77].

The coefficients k_ρ, k_γ and k_β are used for the normalization of ρ_{10}, γ_{10}, β_{10}. The range of stimuli appearing neutral for dark adapted observers is quite large, from 4000K to 11000K and slightly below the black body locus [80]. The equi-energy stimulus, EEW, which lies within this range and also below the black body locus, is mostly used for normalization in CAMs [15]. To obtain an absolute photometric anchor for dark adapted self-luminous stimuli, in addition to the relative colorimetric normalization with respect to the equi-energy stimulus, the coefficients k_ρ, k_γ and k_β were chosen such that all three cone excitations for an equi-energy spectrum are equal to the numerical value of its CIE 1964 10° luminance $L_{10,\mathrm{EEW}}$:

$$\rho_{10,\mathrm{EES}} = \gamma_{10,\mathrm{EES}} = \beta_{10,\mathrm{EES}} = L_{10,\mathrm{EEW}} \tag{5.2}$$

Or

$$k_\rho \int_{390}^{830} \bar{l}_{10}(\lambda)\,d\lambda = k_\gamma \int_{390}^{830} \bar{m}_{10}(\lambda)\,d\lambda = k_\beta \int_{390}^{830} \bar{s}_{10}(\lambda)\,d\lambda = 683.6 \int_{360}^{830} \bar{y}_{10}(\lambda)\,d\lambda \tag{5.3}$$

This yields the following values:

$k_\rho = 666.7$, $k_\gamma = 782.3$ and $k_\beta = 1444.6$.

Using these constants and the absolute spectral radiance of the stimulus, the absolute normalized cone excitations can be calculated from Eq.5.1. Note that in calculating $683.6 \int \bar{y}_{10}(\lambda)\,d\lambda$, changing the integration limits from 360 to 830 nm to 390 to 830 nm does not induce any difference to the constants mentioned above.

5.4.2 Compressed cone responses

A non-linear response compression of the cone excitations [14, 24] is thought to be the first processing step in human vision. It compresses the large optical dynamic range into a rather compact range suitable for encoding. Often, as in CAM97u and CAMFu, this compression is implemented using a sigmoidal curve [12, 14-16, 30]. For a given adaptation state, the intermediate region of this sigmoidal curve (higher than the noise levels and lower than saturation phenomena) can be more or less modelled by a power function. Within the restrictions of the model, i.e. photopic stimuli without glare, the compressed cone responses ρ_c, γ_c and β_c are therefore calculated from the cone excitations ρ_{10}, γ_{10} and β_{10} as follows:

$$\begin{aligned} \rho_c &= \rho_{10}^{c_p} \\ \gamma_c &= \gamma_{10}^{c_p} \\ \beta_c &= \beta_{10}^{c_p} \end{aligned} \tag{5.4}$$

The constant c_p will be determined by fitting the experimental data (see below).

5.4.3 Neural signals

The next stage in colour vision is believed to be a transformation of the compressed responses (Eq.5.4) into three neural signals: the achromatic signal A, and two colour difference signals a and b, respectively related to redness-greenness and yellowness-blueness perception [13, 31]. The achromatic signal is composed of a weighted sum of the three cone responses. The weights were taken in accordance with the estimated numerical distribution of the cones in the retina ρ:γ:β of about 40:20:1 [24, 31, 81, 82]:

$$A = c_A \left(2\rho_c + \gamma_c + \frac{1}{20}\beta_c \right) \tag{5.5}$$

The constant c_A will be determined to obtain an achromatic signal which is correlated to the values calculated by CAM97u.

The colour difference signals a and b are taken to be the same as proposed by Hunt [3] and used in other CAM's [20, 31]:

$$a = c_a\left(\rho_c - \frac{12}{11}\gamma_c + \frac{\beta_c}{11}\right) \tag{5.6}$$

$$b = c_b(\rho_c + \gamma_c - 2\beta_c) \tag{5.7}$$

c_A, c_a and c_b are constants which will be determined by fitting the experimental data (see below).

5.4.4 Hue correlate

It is believed that the ratio of the colour difference signals a and b causes a hue sensation in our visual cortex [13, 31]. By taking the inverse tangent of a and b, the hue angle h can be calculated:

$$h = \frac{180}{\pi}\tan^{-1}(b/a) \tag{5.8}$$

To express hue in terms of a quadrature scale H - i.e. in terms of proportions of the unique hues perceived to be present in the stimulus - the hue angle h is linearly transformed from a 0°-360° range to a 0-400 range:

$$H = H_i + 100\frac{h' - h_i}{h_{i+1} - h_i} \tag{5.9}$$

With h_i the unique hue angle obtained from Hunt [16], H_i the unique hue quadrature, $h' = h + 360$ if h is less than h_1, otherwise $h' = h$, and a value of i chosen so that h' is equal to or greater than h_i and less than h_{i+1} (see Table 5.2).

Table 5.2. Overview of the unique hue data used for calculating the hue quadrature *H [16]*.

Unique hue	Red	Yellow	Green	Blue	Red
i	1	2	3	4	5
h_i	20.14°	90.00°	164.25°	237.53°	380.14°
H_i	0.0	100.0	200.0	300.0	400.0

5.4.5 Colourfulness, brightness, saturation and amount of white correlates

The colourfulness, defined as the perception according to which the perceived colour of an stimulus appears to be more or less chromatic [1], can be represented by the strength of the colour difference signals a and b [13]:

$$M = c_{\mathrm{M}} \times \sqrt{a^2 + b^2} \tag{5.10}$$

With c_{M} a constant to anchor the colourfulness scale of CAM15u to the one used in CAM97u (see below).

A first estimate of the perceived brightness, is given by the achromatic signal A (Eq.5.5) [13]. However, as discussed above, brightness perception is not only dependent on the weighted combination of the cones responses alone but it is also influenced by the strength of the colour of the stimulus (cfr. H-K effect):

$$Q = A + c_{\mathrm{HK1}} \times M^{c_{\mathrm{HK2}}} \tag{5.11}$$

With c_{HK1} and c_{HK2} constant factors used to modulate the strength of the H-K effect and which will be determined by fitting the experimental data (see below).

Analogous to the CIE definition, saturation can be defined as the colourfulness M relative to the brightness Q [1]:

$$s = \frac{M}{Q} \tag{5.12}$$

The amount of white has, as far as we now, never been used and predicted before. From its definition during the experiment, "amount of white" should correlate well to the colourfulness M or saturation s:

$$W = f_{\mathrm{W}}(M \text{ or } s) \tag{5.13}$$

The function f_{W} will be determined by comparing the amount of white perception with the CAM15u saturation and colourfulness.

5.4.6 Determination of the parameters of the model

In addition to the yet to be defined amount of white function f_{W}, the model as proposed in the previous sections has only a few free parameters: c_{p}, c_{A}, c_{a}, c_{b}, c_{M}, c_{HK1} and c_{HK2}.

The parameter c_{p} was determined by optimizing the predictive performance of the model's brightness perception for the largest available set of achromatic stimuli: 15 achromatic stimuli obtained in the 'Random' experiment described above. The brightness of these 15 stimuli, having a 10° FOV and luminance from 5.94 to 297.47 cd/m^2, were rated by a group of 20 observers. For achromatic stimuli, the colourfulness is negligible and the brightness correlate is equal to the achromatic signal A (see Eq.5.5 and Eq.5.11). By minimizing the mean of the squared residual errors between the observed brightness perception and the prediction of the achromatic signal, the optimal value for the parameter c_p was found to be 0.332, which is very close to 1/3. Such a cube root function has often been used to relate physical stimulus quantities to visual sensation: e.g. Leloup [48], Bodmann et al. [83], the CIELAB colour space [70], Schuchard [84], CIECAM02 [12], CAMFu [17]. Instead of this cube root, a log compression has also been adopted by some authors [85]. However, the predictive performance of the CAM15u achromatic signal with a cube root compression was slightly better than the one using a logarithmic compression function (coefficients of determination R^2 were respectively 0.99 and 0.94). Therefore the parameter c_p was fixed to 1/3.

The value of the free parameter c_A (Eq.5.5) was set to 3.22 by anchoring the achromatic signal A of this model to the achromatic signal of CAM97u using the same 15 achromatic stimuli ('Random' experiment). Note that this anchor is limited to the luminance range of these achromatic stimuli, from 5.94 cd/m² to 297.47 cd/m². In Fig.5.2 the achromatic signal of CAM97u for these 15 stimuli is plotted against the one of CAM15u (Eq.5.5). The figure and a coefficient of determination R^2 of 0.99 indicate a good correlation between both.

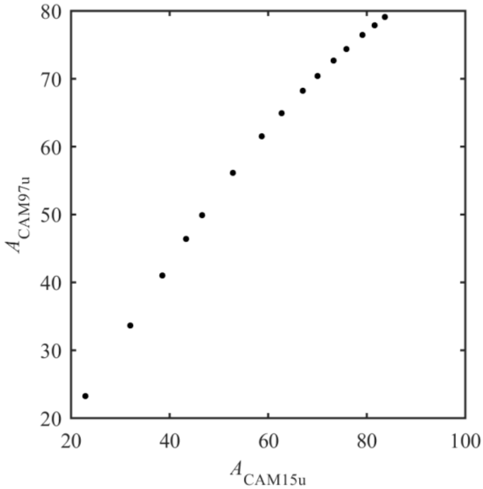


Fig.5.2. Achromatic signal predicted by CAM97u, A_{CAM97u}, versus the one of CAM15u, A_{CAM15u}, for 15 achromatic stimuli of the 'Random' experiment.

The parameters c_a and c_b were determined from the experimental hue quadrature data of the test set by minimizing the mean of the squared residual errors between the experimentally observed hue quadrature H_{avg} and the predicted hue quadrature (Eq.5.9): $c_a = 1$ and $c_b = 0.117$. A correlation coefficient R^2 of 0.99 and a Spearman correlation r of 1.00 (0.996) between the predicted and the observed hue quadrature indicate that Eq.5.9 gives a good prediction of the hue, as illustrated in Fig.5.3. In addition, the goodness-of-fit of the model for predicting the hue, as assessed by the coefficient of variation (CV = 5%), is substantially lower than the inter-observer variability (CV = 10%), indicating the model performs adequately. Considering the unique hue angles, h_i, as free parameters in the model did not substantially improve the hue quadrature prediction ($R^2 = 0.99$).

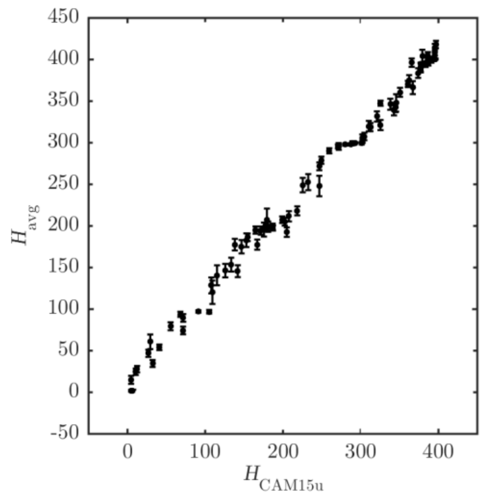


Fig.5.3. Average observed hue H_{avg} with standard error bars versus the hue prediction H_{CAM15u}, for the stimuli of the test set.

Note that Eq.5.9 is a simplified transformation compared to the one used in CAM97u, CAMFu and CIECAM02 in that it eliminates the use of the eccentricity factor. As mentioned in Chapter 2, an eccentricity factor was introduced to compensate for the differences in the strength of perceptual colorization that occurs around the hue circle: for example, the perceptual saturation of a yellow stimulus can never be as high as that of blue stimulus. The eccentricity factor for each unique hue was experimentally obtained in a cone excitation space, resulting in about 0.65 for red, 0.5 for yellow, 1.0 for green and 1.45 for blue [3]. In cube root (compressed) cone response space these eccentricities have respective values of about 0.87, 0.79, 1.00 and 1.13, respectively. In addition, CAM97u takes the Bezold-Brücke effect into account in the hue quadrature equation by making the eccentricity factors of yellow and blue dependent on the luminance of the stimulus [16]. In the CAM15u model the eccentricity factor and consequently the correction for the Bezold-Brücke effect were eliminated as their effect on both colourfulness and hue predictions of the model were found to be negligible.

The parameter c_M was set to 135.52 by anchoring the colourfulness M of the CAM15u model to the colourfulness scale used in CAM97u. In Fig.5.4 the colourfulness of CAM97u for the stimuli of the test set is plotted as a function of the one of CAM15u. The figure and a coefficient of determination R^2 of 0.92 indicate a good correlation between the two.

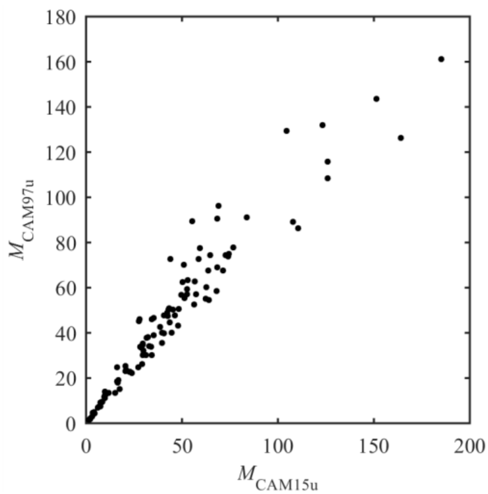


Fig.5.4. Colourfulness predicted by CAM97u, M_{CAM97u}, versus the one of CAM15u, M_{CAM15u}, for the stimuli of the test set.

The parameters c_{HK1} and c_{HK2} (Eq.5.11) were determined by minimizing the mean of the squared residual errors between the experimentally observed and the predicted brightness of the test set. c_{HK1} was found to be equal to 2.559 and c_{HK2} to 0.561. In Fig.5.5 the observed brightness of the stimuli of the test set is plotted against the predicted CAM15u brightness (Eq.5.11). From the figure, a very good correlation between the experiments and the model can be observed which is confirmed by the coefficient of determination R^2 (0.90) and the Spearman correlation r_S (0.95). In addition, the goodness-of-fit of the model's brightness predictor, as assessed by the coefficient of variation (CV = 9%), is substantially lower than the inter-observer variability (CV = 17%).

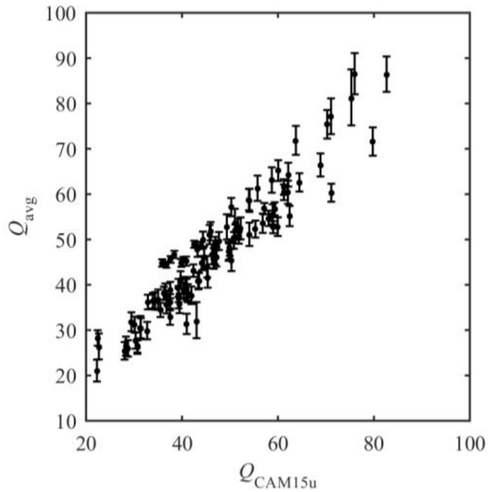


Fig.5.5. 'Average observed' brightness Q_{avg} with standard error bars against the brightness prediction $Q_{CAM15u.}$

A function f_W that predicts the amount of white was obtained by comparing the perceived amount of white of the stimuli of the test set with their CAM15u colourfulness and saturation predictions, see Fig.5.6.

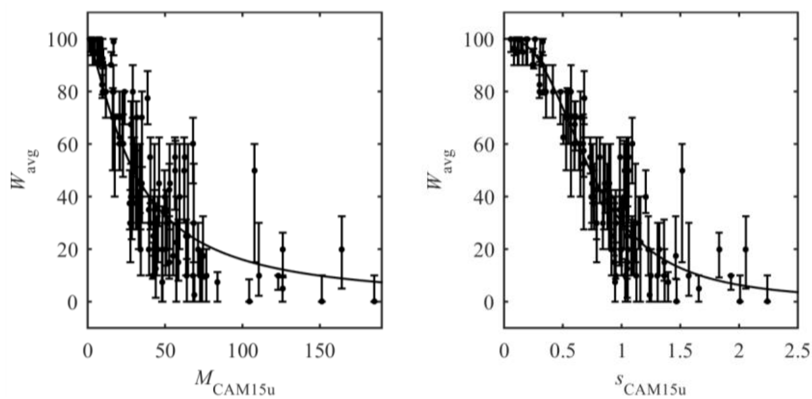


Fig.5.6. 'Average observed' amount of white W_{avg} with interquartile range bars against the CAM15u colourfulness prediction M (left) and the CAM15u saturation prediction s (right).

From the figure it is clear that both colourfulness and saturation exhibit a sigmoidal type relationship with respect to the observed amount of white (full line), with a horizontal asymptote towards 0% white and another one towards 100% white. The large interquartile range bars in the figure indicate the large inter-observer variability of this attribute, as discussed above. The graphs and the values of the Spearman correlation coefficient r_S between the observed amount of white and the predicted CAM15u colourfulness

(r_S = -0.86) and saturation (r_S = -0.90), suggest saturation is the best choice as variable to predict the amount of white. By minimizing the mean of the squared residual errors between the experimentally observed amount of white and a sigmoidal function of the saturation, a prediction of the amount of white is obtained:

$$W = \frac{100}{1 + 2.29 \times s^{2.68}} \tag{5.14}$$

The goodness-of-fit of the model's prediction of the amount of white, as assessed by the coefficient of variation (CV = 23%), is lower than the inter-observer variability (CV = 30%), indicating the model performs adequately. This is also visible in Fig.5.7, where the observed amount of white is plotted against its prediction. The good agreement is also reflected in a high coefficient of determination (R^2 = 0.87) and Spearman correlation coefficient (r_S = 0.90).

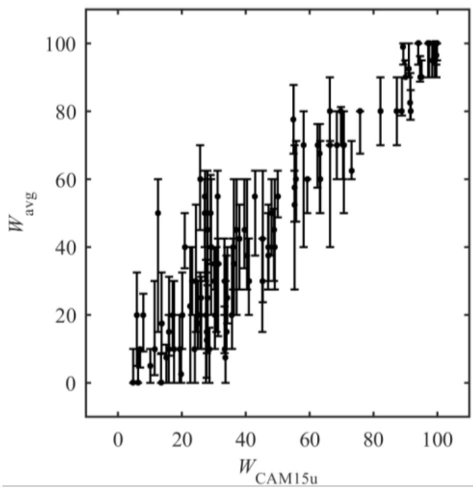


Fig.5.7. 'Average observed' amount of white, W_{avg}, with interquartile range bars against the predicted amount of white, W_{CAM15u} (Eq.5.14).

5.5 Validation

5.5.1 Validation experiment

The performance of the CAM15u model has been verified by the validation set described above and has been compared to that of three other CAMs for unrelated stimuli: CAM97u [16], CAMFu [17] and CAM97um [72]. The model performance was assessed by calculating the coefficient of determination R^2, the Spearman correlation coefficient r_S and the coefficient of variation CV between the mean observer data and those predicted by the models. The model performance indicators for brightness, hue and amount of white are given in Table 5.3. Note that the latter could only be calculated for the CAM15u model.

Table 5.3. Model performance assessed by the coefficient of determination R^2, Spearman correlation coefficient r_S and coefficient of variation CV between the mean observed magnitude of the perceptual attributes obtained in the validation experiments with those predicted by the models.

	Brightness			Hue			Amount of white		
	R^2	r_S	CV	R^2	r_S	CV	R^2	r_S	CV
CAM15u	0.87	0.94	7	0.99	1.00	5	0.76	0.84	32
CAM97u	0.36	0.57	16	0.99	0.99	5	-	-	-
CAM97um	0.80	0.92	9	0.99	0.99	5	-	-	-
CAMFu	0.22	0.41	26	0.99	0.99	5	-	-	-

For brightness, it is clear from the results in Table 5.3 that CAM15u performs best and the model is able to explain 87% of the variance observed in the visual data. The next best model is the modified CAM97um model, which is almost identical to the original CAM97u except that the prediction of brightness has been modified. Surprisingly, the much more simple and direct model CAM15u is able to explain 7% more of the observed variance. The original CAM97u model and CAMFu have a rather low performance: both have low correlation coefficients (R^2 respectively 0.36 and 0.22). The relatively weak performance of these models is confirmed by both the CV values and the graphs in Fig.5.8 where the perceived brightness has been plotted as a function of the model prediction.

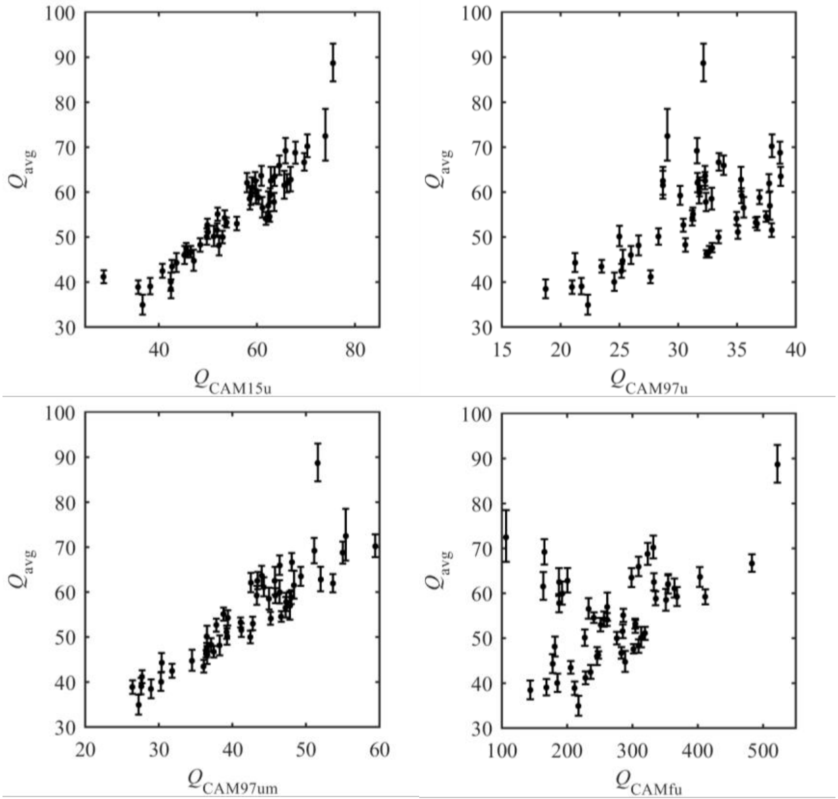


Fig.5.8. 'Average observed' brightness Q_{avg} with standard error bars against the brightness predictions of CAM15u (above, left), CAM97u (above right), CAM97um (below left) and CAMFu (below right) for the unrelated stimuli of the validation set.

For the hue quadrature, all models perform very similar (Table 5.3). All have very high coefficients of variation and Spearman correlation coefficients and the CV values are lower than the inter-observer agreement (11%). The good hue quadrature prediction of all models can also be observed in Fig.5.9. Note that the hue prediction for CAM97um is identical to the prediction of CAM97u.

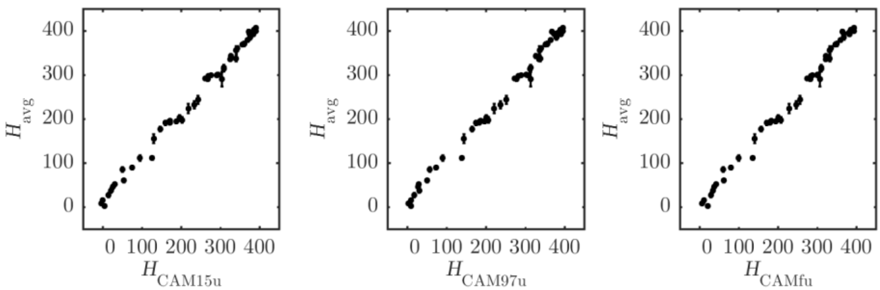


Fig.5.9. 'Average observed' hue quadrature H_{avg} with standard error bars against the hue predictions of CAM15u (left), CAM97u and CAM97um (middle) and CAMFu (right) for the unrelated stimuli of the validation set.

Finally, the observer data for "amount of white" of the validation test set is found to be predicted fairly well by the CAM15u model. Although, the Spearman correlation was not as high as for the brightness and hue predictions, the model still accounted for 76% of the variance in the visual data. In addition, the model prediction CV value (32%) was smaller than the inter-observer CV value (36%). The latter was substantially higher than those for the other attributes, suggesting there was quite a bit of inter-observer disagreement. This can also be observed from the rather large interquartile range bars in Fig.5.10, plotting the perception versus the prediction of this attribute.

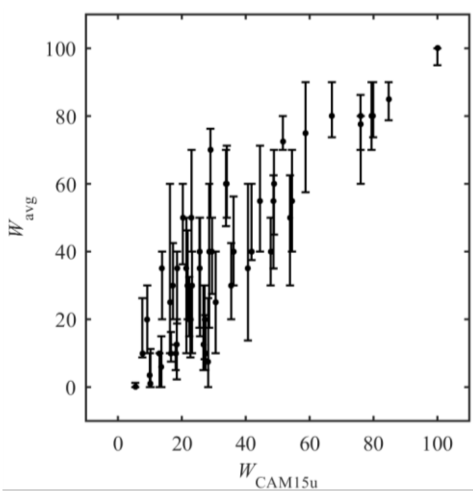


Fig.5.10. 'Average observed' white W_{avg} with interquartile range bars versus the amount of white prediction of CAM15u for the unrelated stimuli of the validation set.

5.5.2 Previous research data

The performance of CAM15u is also validated by comparing the observer perception and the CAM15u prediction of the stimuli used in previous experiments described above: 'Lum6', 'Lum51', 'Match' and 'Random'.

Predictive performance

An overview of the observer variability in terms of the inter-observer CV and the correlation between the observer perceived and CAM15u predicted brightness, hue and amount of white for these four experiments is listed in Table 5.4. For all experiments and all attributes, the CV value of the models prediction against the perception is lower than the inter-observer CV value.

Table 5.4. Overview of the agreement between the 'average observed' perception and the predictions of CAM15u for brightness, hue and amount of white for unrelated stimuli of the 'Lum51', 'Lum6', 'Match' and 'Random' experiment.

		'Lum51'	'Lum6'	'Match'	'Random'
Brightness Q	R^2	0.75	0.75	0.89	0.90
	Spearman r_S	0.87	0.90	0.96	0.94
	CV model	8	10	5	9
	Inter-observer CV	11	13	32	18
Hue H	R^2	0.98	0.99		
	Spearman r_S	0.99	0.99		
	CV model	8	8		
	Inter-observer CV	9	11		
Amount of white W	R^2	0.83	0.89		
	Spearman r_S	0.94	0.94		
	CV model	20	22		
	Inter-observer CV	24	29		

For brightness, the coefficient of determination R^2 is between 0.75 and 0.90 and the Spearman correlation coefficient r_S between 0.87 and 0.96, for the four experiments. As these values are comparable to the values obtained when developing the model, they indicate a good model prediction for brightness. In addition, the 'match' experiment demonstrates that the CAM15u model is also applicable for the prediction of stimuli using a different experimental method. For the 'Random' experiment, which is the most extensive experiment and which was not used in the development of the models, the correlations are higher for CAM15u (R^2 = 0.90 and r_S = 0.94) than for CAM97um (R^2 = 0.81 and r_S = 0.90). These strong validations are also clearly visible in Fig.5.11 where the brightness perception is plotted against the CAM15u prediction. For the 'Match' experiment, the prediction of a matched achromatic reference is plotted against the prediction of the coloured stimuli being matched.

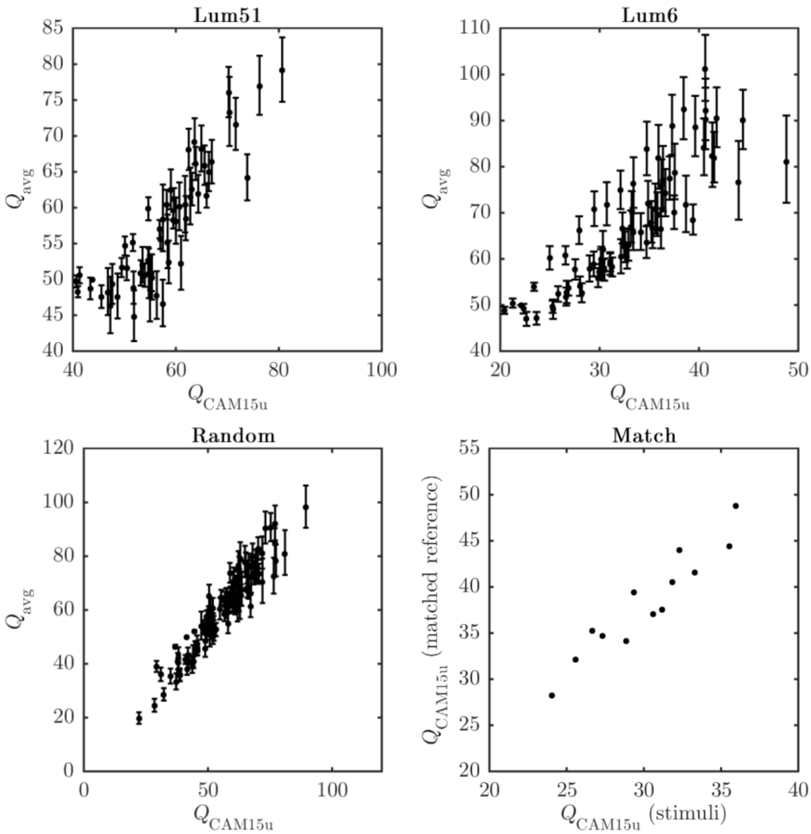


Fig.5.11. 'Average observed' brightness Q_{avg} with standard error bars against the brightness prediction Q_{CAM15u} for the magnitude estimation experiments 'Lum51' (above, left), 'Lum6' (above, right) and 'Random' (below, left). Below, right: prediction of the matched reference against the prediction of the stimuli to be matched for the 'match' experiment.

For hue the correlations are comparable to the ones obtained when developing the model, while for the amount of white the correlations are even higher (Table 5.4); see also Fig.5.12 and Fig.5.13. Note these high correlations for the amount of white for stimuli with a constant luminance, 'Lum51' and 'Lum6', compared to the ones obtained above for stimuli with a variable luminance. This indicates that the amount of white prediction for stimuli with a variable luminance could still be improved by performing new experiments to take the effect of luminance (or brightness) better into account.

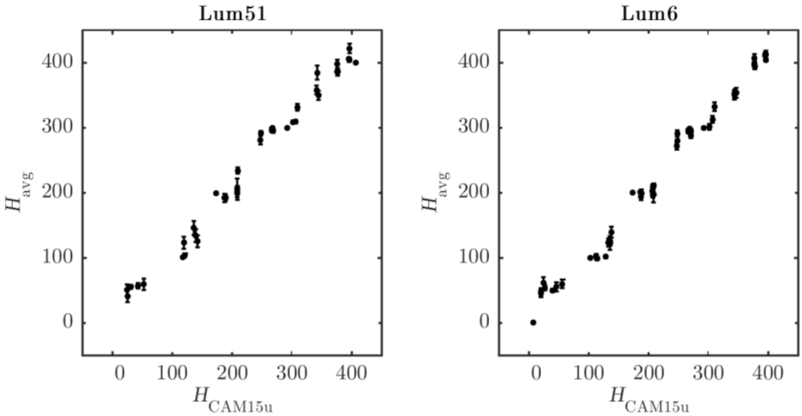


Fig.5.12. 'Average observed' hue H_{avg} with standard error bars against the hue prediction H_{CAM15u} for the experiments 'Lum51' (left) and 'Lum6' (right).

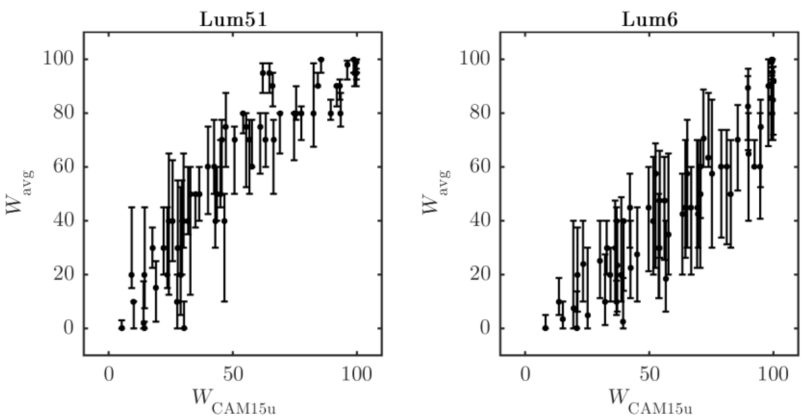


Fig.5.13. 'Average observer' amount of white W_{avg} with interquartile range bars against the amount of white prediction W_{CAM15u} for the experiments 'L51' (left) and 'L6' (right).

5.6 Inconsistencies in the model

Although CAM15u is developed using EEW normalized cone fundamentals and a cube root cone response space, its unique hue angles (Table 5.2) are identical to the ones used in CAM97u, obtained by using cone fundamentals normalized to CIE illuminant C and by using a square root cone response space. Nonetheless, as mentioned before, considering these unique hue angles as free parameters in the model did not substantially improve the hue quadrature prediction.

The redness-greenness a and yellowness-blueness b signals are, except for a weighting factor, similar to the ones used in CAM97u. As mentioned in Chapter 2, these CAM97u a and b signals were obtained by comparing unique hue loci of the NCS scheme with curves of constant ratios of the cone difference signals C_1/C_2. The latter were however calculated using a square root compression and illuminant C normalized cone sensitivities. Although this may cause inconsistencies in CAM15u, these a and b signals have proven their value in other vision models that also use a cube root response space [12, 13]. For the visual data of the test set described above, the use of a more physiologically based redness-greenness, $a = c_{\mathrm{a}}(\rho_{\mathrm{c}} - \gamma_{\mathrm{c}})$, and yellowness-blueness, $b = c_{\mathrm{b}}\left[(\rho_c + \gamma_c)/2 - \beta_c\right]$ correlate, does not provide improvements to the predictive performance of CAM15u. However, an extensive investigation should be performed to study possible improvements induced by using other, physiologically based, expressions for the a and b signals.

5.7 Steps in using CAM15u

Input: Spectral radiance $L_{\mathrm{e},\lambda}(\lambda)$ [$\mathrm{Wnm^{-1}sr^{-1}m^{-2}}$] of the unrelated self-luminous stimulus.

Step 1: Calculate the normalized ρ_{10}, γ_{10} and β_{10} cone excitations directly

$$\begin{aligned}
\rho_{10} &= 666.7\int_{390}^{830} L_{\mathrm{e},\lambda}(\lambda)\bar{l}_{10}(\lambda)\,\mathrm{d}\lambda \\
\gamma_{10} &= 782.3\int_{390}^{830} L_{\mathrm{e},\lambda}(\lambda)\,\bar{\mathrm{m}}_{10}(\lambda)\,\mathrm{d}\lambda \\
\beta_{10} &= 1444.6\int_{390}^{830} L_{\mathrm{e},\lambda}(\lambda)\,\bar{\mathrm{s}}_{10}(\lambda)\,\mathrm{d}\lambda
\end{aligned} \tag{5.15}$$

with $\bar{l}_{10}(\lambda)$, $\bar{\mathrm{m}}_{10}(\lambda)$, $\bar{\mathrm{s}}_{10}(\lambda)$ the CIE 2006 10° cone fundamentals in terms of energy with a 1 nm spacing, available on the website http://www.cvrl.ac.uk. When the radiance is not available, the absolute 10° tristimulus values X_{10}, Y_{10}, Z_{10} of the stimulus can be used as input. Step 1 is then replaced by a direct conversion of these tristimulus values into an approximation of the normalized cone excitations ρ_{10}, γ_{10}, β_{10} (see below, Section 5.9).

Step 2: Calculate the compressed cone responses by taking the cube root of the cone excitations

$$\begin{aligned} \rho_c &= \rho_{10}^{1/3} \\ \gamma_c &= \gamma_{10}^{1/3} \\ \beta_c &= \beta_{10}^{1/3} \end{aligned} \tag{5.16}$$

Step 3: Calculate the achromatic signal and the colour difference signals

$$A = 3.22\left(2\rho_c + \gamma_c + \frac{1}{20}\beta_c\right) \tag{5.17}$$

$$a = \rho_c - \frac{12}{11}\gamma_c + \frac{\beta_c}{11} \tag{5.18}$$

$$b = 0.117(\rho_c + \gamma_c - 2\beta_c) \tag{5.19}$$

Step 4: Calculate the hue angle and hue quadrature

$$h = \frac{180}{\pi}\tan^{-1}(b / a) \tag{5.20}$$

$$H = H_i + 100\frac{h' - h_i}{h_{i+1} - h_i} \tag{5.21}$$

With $h' = h + 360$ if h is less than h_1, otherwise $h' = h$, and a value of i chosen so that h' is equal to or greater than h_i and less than h_{i+1}. With h_i and H_i equal to:

Unique hue	Red	Yellow	Green	Blue	Red
i	1	2	3	4	5
h_i	20.14°	90.00°	164.25°	237.53°	380.14°
H_i	0.0	100.0	200.0	300.0	400.0

Step 5: Calculate the colourfulness, brightness and saturation

$$M = 135.52 \times \sqrt{a^2 + b^2} \tag{5.22}$$

$$Q = A + 2.559 \times M^{0.561} \tag{5.23}$$

$$s = \frac{M}{Q} \tag{5.24}$$

Step 6: Calculate the amount of white

$$W = \frac{100}{1 + 2.29 \times s^{2.68}} \tag{5.25}$$

5.8 Worked example

The CAM15u model gives the following results for a 30.00 cd/m² sample with a spectral radiance given in Fig.5.14: $\rho_{10} = 29.36$, $\gamma_{10} = 33.07$, $\beta_{10} = 38.06$, $\rho_c = 3.09$, $\gamma_c = 3.21$, $\beta_c = 3.36$, $A = 30.75$, $a =$ -0.11, $b =$ -0.05, $h = 204.57$, $H = 255.02$, $M = 16.49$, $Q = 43.07$, $s = 0.38$, $W = 85.13$.

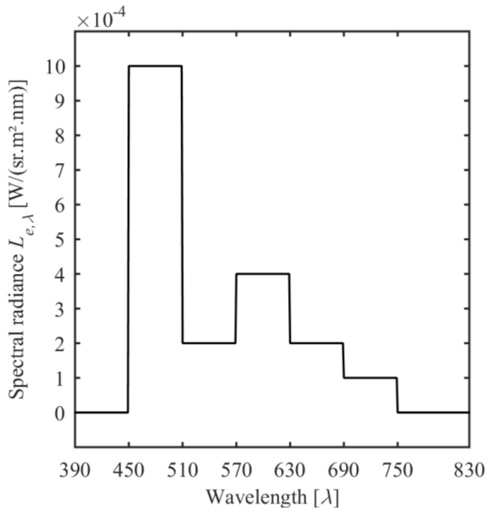


Fig.5.14. Spectral radiance of the sample used in the worked example.

5.9 Conversion from tristimulus values into cone excitations

When the spectral radiance of the stimulus is not available but the absolute tristimulus values X_{10}, Y_{10}, Z_{10} are, the normalized cone excitations can be approximated as:

$$\begin{bmatrix} \rho_{10} \\ \gamma_{10} \\ \beta_{10} \end{bmatrix} = \begin{bmatrix} 0.211831 & 0.815789 & -0.042472 \\ -0.492493 & 1.378921 & 0.098745 \\ 0 & 0 & 0.985188 \end{bmatrix} \begin{bmatrix} X_{10} \\ Y_{10} \\ Z_{10} \end{bmatrix} \tag{5.26}$$

For the worked example, a mean difference of 0.0003% and 3.05% was found between the attributes calculated according to the cone fundamentals and the attributes obtained using Eq.5.26, whereby the *XYZ* values have been respectively calculated with the new CIE 2006 *XYZ* and the CIE 1964 *XYZ* CMF's.

5.10 Brightness scale

To estimate the brightness scale of CAM15u, in Table 5.5, the chromaticity values, luminance and CAM15u brightness prediction of some stimuli are listed. The data are calculated from the spectrum of five stimuli used in the visual tests of the doctoral research project. In Fig.5.15 the luminance values and brightness predictions are plotted.

Table 5.5. Overview of the chromaticity coordinates, luminance value and CAM15u brightness prediction of some unrelated self-luminous stimuli.

Colour	u'_{10}	v'_{10}	L_{10} (cd/m^2)	Q_{CAM15u}
Red	0.4734	0.5270	6	48.00
			30	70.90
			50	80.46
Green	0.0877	0.5684	6	40.33
			30	60.85
			50	69.54
Yellow	0.3402	0.5405	6	38.24
			30	57.92
			50	66.27
Blue	0.1263	0.2526	6	44.01
			30	66.05
			50	75.35
White	0.2339	0.4899	6	25.65
			30	41.03
			50	47.74
			100	58.73
			150	66.35
			200	72.38
			250	77.44
			300	81.85

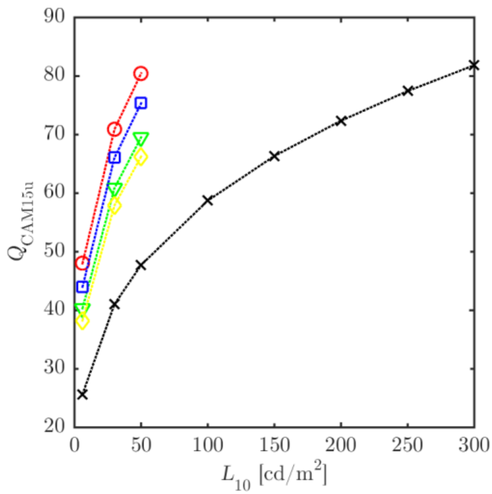


Fig.5.15. Luminance value and corresponding CAM15u brightness prediction of some unrelated self-luminous stimuli.

5.11 Conclusions

The brightness, hue and "amount of white" perception of a set of unrelated self-luminous stimuli was investigated in a magnitude estimation experiment with twenty observers. The amount of white is a new attribute, and basically corresponds to a layperson's conception of attributes such as colourfulness, chroma or saturation. It was introduced based on the results of a pilot study that showed that laypersons often have difficulty to understand and hence to judge the colourfulness of a stimulus in an experiment. A non-forced hue evaluation method revealed that the hue perception of a substantial part of the observers, 20%, cannot be mapped to a hue quadrature scale, commonly believed to be representative of typical hue perception of observers.

Based on the obtained visual data, a new colour appearance model for unrelated self-luminous stimuli, CAM15u, was developed. The main features of the model are the use of the absolute spectral radiance of the stimulus as input, the use of the CIE 2006 cone fundamentals and a simplified calculation procedure compared to existing models. The model predicts the brightness, hue, colourfulness, saturation and the amount of white. The CAM15u model is restricted to photopic, non-glare-inducing unrelated stimuli having a field of view of 10°.

An additional magnitude estimation experiment was carried out to validate the CAM15u model and to compare its predictive performance with that of other CAMs for unrelated colours like CAM97u, CAM97um and CAMFu. It was found that, despite its simplicity, CAM15u performs better or at least equally well compared to the existing CAMs. The excellent performance of CAM15u was also confirmed by comparing its predictions with the visual data obtained in Chapter 4 (both magnitude estimation and matching experiments).

Chapter 6

SIZE EFFECT ON BRIGHTNESS

Like many people, I was facing a renovation project after buying our house. Some structural problems required my attention: the steep stairs, the roof without isolation, etc. Meanwhile, my wife was making our house a home by personalizing the rooms. According to my wife, one of the walls, excellently built and painted in good quality orange paint, really needed to be taken care of. After realising that the hue was the problem, I found myself struggling with test colour patches of about 5 on 10 cm. I chose all kind of hues having the same (very high) saturation, from green to red, yellow to blue. After losing the discussion about saturation, I accompanied my wife to the shop and bought a bright (the only thing we agreed on) … greyish white paint. Later that day, we painted the wall but found the result deviating a lot from our expectations. Surprised by the fact that the appearance of a small patch was quite different compared to the one of the whole wall, we worked till late that night and found our greyish white wall under tungsten light…

The colour appearance of a stimulus depends on the size of that stimulus. In this chapter, the effect of stimulus size on the brightness perception is investigated. The CAM15u model is extended to CAM15us, being able to predict the brightness of different size stimuli.

6.1 Introduction

As discussed in Chapter 4, the brightness perception of a stimulus cannot be predicted using a luminance based approach alone [86]. A more complex approach including among others cone-compression, opponent modulation, non-linearity of the human visual system and the Helmholtz-Kohlrausch effect is needed, e.g. by using a CAM. In the previous chapter, CAM15u was developed and found to perform better compared to other CAMs for unrelated colours [76]. The new CAM15u is applicable to photopic, non-glare-inducing stimuli with a fixed field of view (FOV) of 10°. However, the size of the stimulus has been shown to substantially affect colour perception: the larger the stimulus, the higher the brightness [87], lightness, chroma and colourfulness [17, 88, 89].

This chapter deals with the effect of stimulus size on brightness [90]. The effect was investigated in a series of magnitude estimation experiments in which twenty observers had to rate the brightness of unrelated self-luminous circular stimuli with FOVs between 1° and 30° (1°, 2°, 5°, 10°, 15°, 20°, 25° and 30°). Based on the visual data, (only) the brightness prediction of the CAM15u model was extended to include the effect of stimulus size. The extended model, referred to as CAM15us, has also been validated using additional visual data.

6.2 Existing models predicting the size effect on brightness

Ronchi [91] investigated the brightness of white circular stimuli (x, y = 0.495, 0.414) with the same luminance (36 cd/m²) but increasing size (from 1° to 3° FOV). During the experiment, a reference stimulus of 1° and a test stimulus of variable size were viewed simultaneously on a grey background (13.50 cd/m²) at a distance of 59 cm. The ratio of the brightness of the test stimulus to that of the reference one, was evaluated by 10 observers using the magnitude estimation method. The following ratio of relative brightness, R, was found:

$$R = \frac{Q_{\mathrm{t}}}{Q_{\mathrm{r}}} = (A_{\mathrm{t}} - A_{\mathrm{r}})e^{0.54} \tag{6.1}$$

with Q the brightness and A the area of the test and reference stimuli, respectively denoted by the subindices t an r.

Gombos and Schanda [92] have also investigated the effect of size on brightness perception for circular stimuli with an identical luminance. Two stimuli of different sizes were shown simultaneously for a few seconds and observers were asked to scale their brightness. Stimuli of 1° FOV seemed to have a 20% to 30% lower brightness than disks of 3° FOV or larger.

Xiao [89, 93-95] has investigated the effect of size on lightness, chroma and hue perception, however not on brightness. Two models were developed to predict the change in colour appearance from a 2° stimulus to a larger sized stimulus: the size effect *transform* and the size effect *correction*. The size effect transform can be used in the first stage of a colour appearance model by transforming the *LMS* tristimulus values from one stimulus size to another. The size effect correction model provides size dependent CIECAM02 lightness, chroma and hue quadrature attributes. The performances of Xiao's models were verified and compared using experimental data obtained from a colour matching experiment in which ten observers evaluated the lightness, chroma and hue of ten coloured square stimuli presented in six different sizes (2°, 8°, 19°, 22°, 44° and 50°). The results showed that the size effect transform performed better than the size effect correction [94]. The following size effect transformation was derived:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix}_{\theta} = \begin{bmatrix} 1.306 & 0.328 & 0.193 \\ -0.632 & 2.176 & 0.274 \\ -0.543 & 0.047 & 2.241 \end{bmatrix} \begin{bmatrix} \alpha(\theta) & 0 & 0 \\ 0 & \beta(\theta) & 0 \\ 0 & 0 & \gamma(\theta) \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}_{2^\circ} \tag{6.2}$$

Where $\alpha(\theta)$, $\beta(\theta)$, $\gamma(\theta)$ represent the changes in cone responses between a target stimulus size of θ (expressed in degrees) with respect to the 2° values:

$$\begin{aligned} \alpha(\theta) &= 0.000062\theta^2 + 0.00580\theta + 0.5106 \\ \beta(\theta) &= 0.000064\theta^2 + 0.00556\theta + 0.5154 \\ \gamma(\theta) &= 0.000090\theta^2 + 0.00280\theta + 0.5184 \end{aligned} \tag{6.3}$$

Later on in this chapter, the performance of these three models will be evaluated using the data collected in a series of visual experiments.

Finally, as mentioned in Chapter 2, Fu et al. [17] have investigated the size effect of unrelated colours on brightness, colourfulness and hue perception. In a magnitude estimation experiment, ten observers evaluated the colour appearance of circular stimuli having a FOV of 0.5° and 10° under photopic and 0.5°, 1°, 2° and 10° under mesopic viewing conditions. A colour appearance model, CAMFu, predicting the visual attributes of unrelated colours was developed based on the results of the experiments [17]. The performance of this model has already been examined in Chapter 4, wherein it was shown that Fu's model substantially underestimates the Helmholtz-Kohlrausch effect and is thus unable to accurately predict the brightness of coloured stimuli.

6.3 Experimental setup and method

Observers were asked to evaluate the perceived brightness of stimuli using the magnitude estimation method described in Chapter 3. The experiments were carried out in a darkened viewing room and the stimuli were presented on the LCD monitor (see Chapter 3). On this monitor, circular stimuli with a FOV of 1° to 30° were presented to observers seated in front of the monitor and leaning on a fixed chinrest. All stimuli were spectrally characterized before and after the experimental campaign, which lasted for about one month. With a mean $\Delta E_{u'v'}$ of 0.0026 and a maximum $\Delta E_{u'v'}$ of 0.0049, the colour differences between all pairs of stimuli before and after the experiment campaign are acceptable [96]. The variability in 10° luminance settings between pairs, was found not to exceed 2% (mean 0.61%) of the pair's mean luminance.

Based on the studies of Fotios and Cheal [97], Loe et al. [98], Dubois [99], Cowdroy [100] and Lau [101], the CIE [45] proposed - as a first estimate - that a centrally fixated visual field larger than 20° adequately represents the spatial brightness response of larger fields, including full field vision. To verify the CIE proposal, our experiments were carried out with even larger stimuli. Eight different sized circular stimuli were chosen for this experiment: 1°, 2°, 5°, 10°, 15°, 20°, 25° and 30°. These stimuli were viewed from a distance of approximately 42 cm and have a diameter of approximately 0.7, 1.5, 3.7, 7.3, 11.0, 14.7, 18.5 and 22.4 cm, respectively. The luminance of the stimuli varied between 5.95 and 199.42 cd/m². All stimuli were presented against a black background consisting of part of the LCD monitor, with a FOV up to 42°x64° and a 10° luminance between 0.2 and 0.5 cd/m², and the black walls with a luminance below 0.2 cd/m².

6.4 Visual tests

The effect of size on brightness perception of unrelated self-luminous stimuli presented on an LCD monitor has been investigated in a series of psychophysical experiments using three different sets of stimuli. A first set, referred to as the '*CAM15u validation set*', was used to confirm the applicability of the CAM15u brightness prediction for the experimental setup using an LCD monitor. The set was composed of 10° stimuli and was evaluated by eight observers. Then, to investigate the size effect on brightness, twenty observers evaluated two series of different sized stimuli: a '*general test set*' to investigate the size effect on brightness and to develop a size dependent brightness prediction and a '*general validation set*' to validate this new prediction.

6.4.1 CAM15u validation set

Eight observers (5 male and 3 female) with ages ranging between 21 and 28 years (average 25) participated in the experiment validating the CAM15u brightness prediction of 10° stimuli. All observers had normal colour vision according to the Ishihara 24 plate Test for Colour Blindness and had participated in previous experiments. The observers were asked to rate 55 test stimuli with a FOV of 10° in comparison with a reference 10° achromatic stimulus. The luminance of the test stimuli ranged between 5.85 and 153.36 cd/m² and their chromaticity coordinates are plotted in Fig.6.1. The luminance of the reference stimulus was 43.32 cd/m² with a chromaticity close to that of illuminant D65 (u'_{10}, v'_{10} = 0.1979, 0.4695; $\Delta E_{u'v'}$ = 0.0027).

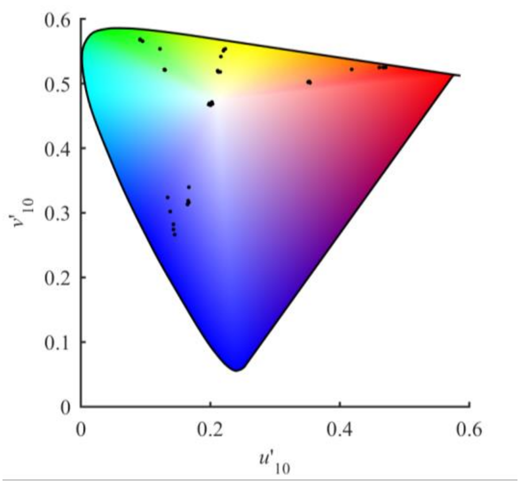


Fig.6.1. CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the 55 stimuli of the 'CAM15u validation set'.

The inter- and intra-observer variability was assessed using the coefficient of variation (CV), Eq.3.1. The good average inter-observer CV value, 17%, as well as the small CV range (from 10% to 24%), indicate observers agreed well and had little difficulties in scaling brightness. The mean CV value is comparable to the values reported in Chapters 4 and 5 (8%, 13%, 14%, 17% and 18%) and better than the ones reported by Fu et al. [17] (29 %) and Koo and Kwak [69] (40%). The average CV value for intra-observer variability was calculated to be 13%, which is in line with the 11%, 12% and 20% reported in Chapters 4 and 5 and the 15% repeatability obtained by Fu et. al [17].

6.4.2 General test set

With the aim of extending the CAM15u brightness prediction to include the effect of stimulus size, twenty observers, 9 male and 11 female - with ages ranging between 20 and 32 years (average 25, median 24), were asked to rate the brightness of different sized test stimuli with respect to the brightness of a 10° identically coloured reference stimulus. All observers had normal colour vision according to the Ishihara 24 plate Test for Colour Blindness. Seventeen of the observers had participated in previous experiments while the other three were naïve with respect to the purpose of the experiment. The stimulus presentation sequence was as follows. First, twenty red stimuli, composed of 12 red 'warming up' and 8 red 'test' stimuli, were rated in comparison with a 10° red stimulus as reference. The 'test' stimuli each had a different, randomly ordered size of 1°, 2°, 5°, 10°, 15°, 20°, 25° and 30° FOV, while the 'warming up' stimuli had sizes randomly selected from these eight FOVs. After a break of 30 seconds twenty blue stimuli were evaluated against a 10° blue reference. Subsequently, yellow, green and white stimuli were evaluated in a similar manner. The 10° luminance (cfr. CIE 1964 observer) and chromaticity coordinates of the stimuli are given in Table 6.1.

Table 6.1. Luminance values and chromaticity coordinates of the stimuli of the 'general test set'.

	L_{10} (cd/m²)	u'_{10}	v'_{10}
Red	20.00	0.4571	0.5239
Blue	22.33	0.1446	0.2686
Yellow	19.86	0.2165	0.5513
Green	20.35	0.0967	0.5654
Achromatic	102.23	0.1958	0.4690

As before, inter-observer variability was assessed by the coefficient of variation. The average inter-observer CV value, 21%, as well as the CV range (from 8% to 50%), indicate observers agreed slightly less compared to the brightness evaluation of stimuli having the same size (see Chapters 4 and 5).

6.4.3 General validation set

Finally, visual data were collected to validate the CAM15us size dependent brightness prediction. In the experiment, the same twenty observers as in the 'general test set' were asked to evaluate a set of 'general validation stimuli' composed of 19 achromatic and 43 coloured stimuli with respect to a reference 10° achromatic stimulus. The luminance of the 19 achromatic stimuli ranged from 5.95 to 199.42 cd/m², with chromaticity coordinates close to that of illuminant D65 (u'_{10}, v'_{10} = 0.1979, 0.4695; mean $\Delta E_{u'v'}$ = 0.0022) and a FOV of 10°. The luminance of the achromatic reference stimulus was 102.14 cd/m², approximately in the middle of the range of the achromatic test stimuli. The luminance of the 43 random coloured stimuli varied between 6.01 and 50.16 cd/m², having random chromaticity coordinates and sizes between 1° and 30° (see Fig.6.2). The mean CV values for the inter- and intra-observer agreement, respectively 20% and 16%, are in line with those mentioned earlier.

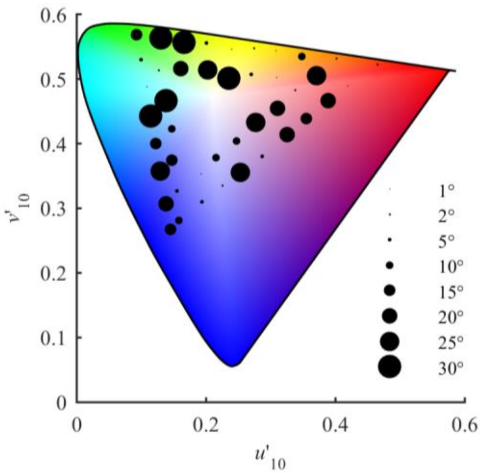


Fig.6.2. CIE 1976 u'_{10}, v'_{10} chromaticity coordinates of the 43 coloured stimuli of the 'general validation set', highlighted according to their size.

6.5 Results

6.5.1 CAM15u validation

CAM15u, developed based on visual data obtained using the LED module surrounded by a dark wall, has been applied to the stimuli presented on the display. Despite the differences in experimental setup, the model's brightness predictor, Q_{CAM15u}, was able to explain 79% of the variation in the *CAM15u* *validation set*, which is only slightly lower than the value reported on the validation set using the LED module (R^2 = 87%). This coefficient of determination, as well as the Spearman correlation ($r_S = 90$) and goodness-of-fit as assessed by the coefficient of variation (CV = 14%, inter-observer CV = 17%) between the CAM15u brightness prediction and the average observer data, confirm the model's adequate performance under the new experimental conditions. The CAM15u brightness predictions versus the observer data are illustrated in Fig.6.3.

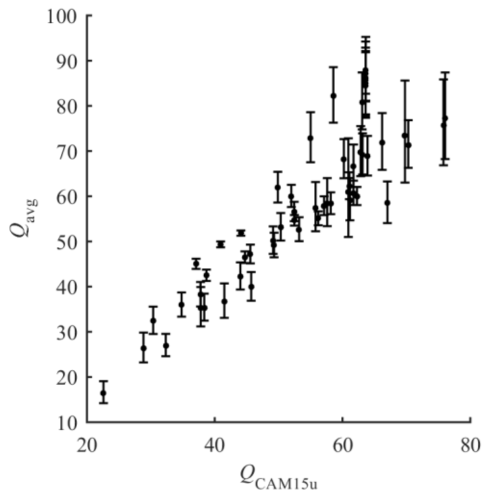


Fig.6.3. 'Average observer' brightness (Q_{avg}) with standard error bars plotted against the CAM15u brightness predictions for the stimuli of the 'CAM15u validation set'.

6.5.2 Size dependent brightness prediction, CAM15us

The effect of size on perceived brightness for white and four different hues is illustrated in Fig.6.4, in which the average observer brightness (Q_{avg}) is plotted as a function of the FOV of the stimuli of the '*general test set*'. The very similar observer responses for all the chromaticities suggest the size effect to be hue independent. It was found that the effect of stimulus size (in terms of FOV expressed in degrees) on brightness, relative to the 10° reference, $Q_{10,ref}$, could be modelled very well using a single power function:

$$\frac{Q_{avg}}{Q_{10,ref}} = \left(\frac{\mathrm{FOV}}{10°}\right)^{0.271} \tag{6.4}$$

The coefficient of determination and the coefficient of variation between the brightness calculated using Eq.6.4 and the average observer data were respectively R^2 = 0.95 and CV = 6%. Note that the latter was also substantially lower than the inter-observer variability (CV = 21%) further confirming the excellent goodness-of-fit of the model. Finally, the hue independence was confirmed by the low CV values, i.e. 4%, 4%, 4%, 3% and 10%, for the red, yellow, green, blue, and white stimuli respectively.

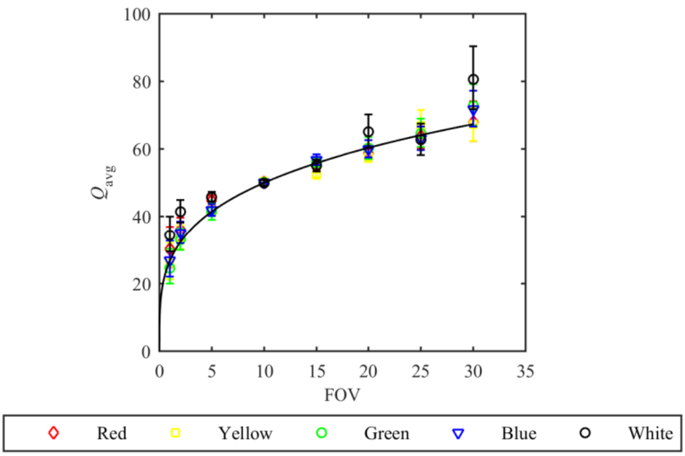


Fig.6.4. Average observer brightness (Q_{avg}) with standard error bars as a function of the FOV of the stimuli of the 'general test set'. The modelled brightness prediction (Eq.6.4) is also plotted (black line).

As the brightness of 10° stimuli is well predicted by the CAM15u model, Eq.6.4 can be easily re-formulized to a ***size dependent*** CAM15us brightness prediction, Q_{CAM15us}, as follows:

$$Q_{\text{CAM15us}} = Q_{\text{CAM15u}}\left(\frac{\text{FOV}}{10}\right)^{0.271} \tag{6.5}$$

As mentioned earlier, Ronchi [91] used the stimulus area to predict the brightness of white stimuli from 1° to 3° FOV. The predictions of Ronchi's brightness model were examined for the 1° and 2° stimuli of the 'general test set' by comparing the relative brightness ratio as calculated using Eq.6.1 to the values obtained from the visual data. It was found that the brightness ratios calculated for the red, yellow, blue, green and white stimuli (resp. 1.19, 1.29, 1.31, 1.34, 1.21) were substantially smaller than the theoretical brightness ratio value of 2.15 as found using Eq.6.1. The disagreement could be due to Ronchi's use of a non-dark 13.5 cd/m² background or due to her model using absolute areas without taking the viewing distance into account as in the FOV. Thereby the validity of her model is limited to the 13.5 cd/m² grey background and 59 cm viewing distance adopted in her experiments.

Gombos and Schanda [92] found the brightness of a 1° stimulus to be 20% to 30% lower than an identical stimulus with a 3° FOV. In the general test set, as described above, the brightness appearance of the 1° red, yellow, blue, green and white stimulus was found to be respectively 16%, 22%, 23%, 26% and 17% lower than the same stimulus having a 2° FOV, in general agreement with the results of Gombos and Schanda using a 3° FOV.

As mentioned earlier, the CIE proposed that the brightness of a centrally fixated visual field larger than 20° adequately represents the spatial brightness response of larger fields, including full field vision. However, the data plotted in Fig. 6.4. suggest there is still a small increase in brightness with increasing field size beyond 20°. Further research with field sizes up to more than 30° is necessary to find the field size adequately representing the spatial brightness response of full field vision.

6.5.3 CAM15us validation

The performance of the CAM15us brightness prediction (Eq.6.5) has been verified using the results of the 'general validation set' described above. In Fig.6.5 the 'average observer' brightness (Q_{avg}) for the stimuli of the 'CAM15us validation set' is plotted against the size dependent brightness prediction Q_{CAM15us} (left). From this figure, it is clear that the size dependent CAM15us (Eq.6.5) is a very good predictor as indicated by the high coefficient of determination, R^2 (0.96) and high Spearman correlation r_{S} (0.98). The goodness-of-fit of the Q_{CAM15us} prediction as assessed by the coefficient of variation (CV = 8%) is also much lower than the inter-observer variability (CV = 20%), indicating that the CAM15us model performs adequately. The need for a size dependent brightness prediction is clearly shown by comparing the model performance in predicting the brightness of the 'general validation set' of CAM15us with that of CAM15u (R^2= 0.44, Spearman r = 0.64, CV = 20%).

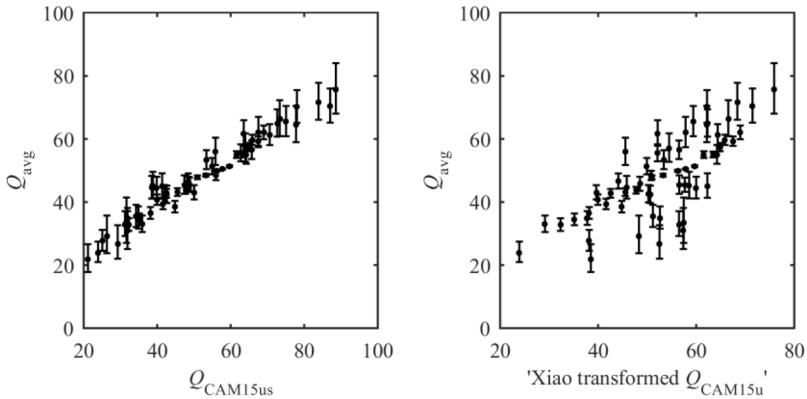


Fig.6.5. Average observer brightness (Q_{avg}) with standard error bars against the size dependent CAM15us brightness prediction using Eq.6.5 (left) and the brightness prediction based on the Xiao size effect transformation included in CAM15u (right) for the stimuli of the 'CAM15us validation set'.

It could be interesting to investigate the use of Xiao's size effect transform (Eq.6.2-6.3) in CAM15u as an alternative to the power function used in CAM15us (Eq.6.5). As the size effect transform normally uses 2° *LMS* values as input while CAM15u was developed for a 10° *LMS* input, the size effect transform was first modified to account for 10° *LMS* values as input. The result of this 'Xiao transformed CAM15u' prediction of brightness is shown in Fig.6.5 (right). With R^2 and Spearman r_S values being equal to respectively 0.57 and 0.75, it is obvious that this procedure leads to a rather weak correlation. One of the reasons could be that Xiao's model was constructed for *LMS* values based on the CIE 1931 colour matching functions. In addition, there is a logical inconsistency in the Xiao transform as there is no identity transformation between the 2° *LMS* and itself ($\theta = 2°$ in Eq.6.2).

6.6 Conclusions

The brightness perception of different sized, unrelated self-luminous stimuli was investigated in a series of magnitude estimation experiments. A substantial, hue independent, effect of stimulus size on brightness was found. The impact of stimulus size on brightness was incorporated into the brightness prediction of the CAM15u model. The size effect could be effectively modelled by a simple power function. The predictive performance of the modified brightness prediction, $Q_{CAM15us}$, was validated using the results obtained in an additional magnitude estimation experiment in which twenty observers evaluated the brightness of unrelated self-luminous stimuli having variable size, chromaticity and luminance. Finally, the performance of the size effect transform as proposed by Xiao was found to be inferior to the performance of CAM15us.

Chapter 7

CONCLUSIONS

In the last year of my doctoral research, I finally felt PhD worthy, as in my opinion, some big steps were taken in the development of a new colour appearance model and, not unimportantly, in my personal 'development'. All of the work and gathered knowledge came together in a very interesting final '100 meters'. Proudly, I can look back at my accomplishments and at the time spent at the lab. Now, new challenges are awaiting…

In this chapter the main conclusions, some applications and future research possibilities are summarized.

7.1 Conclusions

The failure of CAM97u [16], CAMFu [17] and other vision models in predicting the brightness of unrelated stimuli, has narrowed the aim of the doctoral research project from its original goal of developing a general new CAM for self-luminous stimuli to developing a new CAM for *unrelated* self-luminous stimuli (stimuli viewed on a black background). In a series of psychophysical experiments, the appearance of these unrelated self-luminous stimuli was investigated. The results of these visual experiments were used to develop a colour appearance model for unrelated self-luminous stimuli, CAM15u, as well as an extension (CAM15us) that includes the effect of stimulus size on brightness. The models can be used to improve other CAMs and can be extended to other viewing conditions, e.g. neutral or coloured self-luminous backgrounds.

7.1.1 CAM97um

In a first extensive psychophysical experiment, the brightness of a set of 58 unrelated self-luminous coloured stimuli with a FOV of 10° and a luminance of 51.37 cd/m^2 was investigated in a magnitude estimation experiment with nine observers [23]. The correlation between the brightness perception of these observers and the brightness calculated according to the models based on the equivalent luminance ($L_{\mathrm{Eq,CIE}}$ [53] and $L_{\mathrm{Eq,Nay}}$ [52]), the CAM97u model [16], the ATD01 model [63] and the CAMFu [17] model has been investigated. Although the models included the H-K effect and half of the models were developed to work with unrelated colours, none of the models seemed to be able to adequately predict the perceived brightness.

In a second series of psychophysical experiments, the brightness of a set of 58 unrelated self-luminous coloured stimuli with a 10° FOV and a constant luminance of 6.23 cd/m^2, and of a set of 17 achromatic 10° stimuli, with luminance values ranging from 7.54 cd/m^2 to 47.60 cd/m^2, was investigated in a magnitude estimation experiment with twenty observers [72]. It was found that the H-K effect contributed significantly to the observed brightness. The brightness prediction of the existing vision models was investigated but, again, none of the models performed satisfactorily. Adapting the CAM97u model by increasing the colourfulness contribution in the brightness attribute resulted in a modified model, called CAM97um, which allows for a substantially better brightness prediction.

The performance of the new model was confirmed by both a matching experiment with 13 unrelated self-luminous coloured 10° stimuli and an extensive validation magnitude estimation experiment using a random sequence of 107 stimuli with a FOV of 10°, a wide chromaticity range, and a wide range of luminance values. The modified model CAM97um clearly outperformed the other existing vision models and was found to give a reliable brightness prediction for unrelated self-luminous stimuli.

7.1.2 CAM15u

In a third series of psychophysical experiments, the brightness, hue and "amount of white" perception of a set of 105 unrelated self-luminous stimuli with 10° FOV, with luminance values ranging from 6 cd/m^2 to 60 cd/m^2 and with a wide chromaticity range, was investigated in a magnitude estimation experiment with twenty observers [76]. The amount of white is a new attribute, and basically corresponds to a layperson's conception of attributes such as colourfulness, chroma or saturation. It was introduced after a preliminary pilot study indicating that laypersons often have difficulty understanding and hence judging colourfulness of a stimulus. Although the amount of white may be a more familiar attribute than the colourfulness, unfortunately it generally did not lead to a more robust estimate. This is probably a result of the increased difficulty of quantifying the amount of white as the stimulus becomes more saturated. However, because of its familiarity and simplicity, amount of white has been used throughout the experiments. Furthermore, a non-forced hue evaluation method revealed that the hue perception of a substantial part of the observers, 20%, could not be mapped to a hue quadrature scale, commonly believed to be representative of typical hue perception of observers.

Based on the obtained visual data, a new colour appearance model for unrelated self-luminous stimuli, CAM15u, was developed. The main features of the model are the use of the absolute spectral radiance of the stimulus as input, the use of the CIE 2006 cone fundamentals and a simplified calculation procedure compared to existing models. The model predicts the brightness, hue, colourfulness, saturation and the amount of white. The CAM15u model is restricted to photopic, non-glare-inducing unrelated stimuli having a field of view of 10°.

An additional magnitude estimation experiment with 52 stimuli was carried out to validate the CAM15u model and to compare its predictive performance with that of other CAMs for unrelated colours like CAM97u, CAM97um and CAMFu. It was found that, despite its simplicity, CAM15u performs better or at least equally well.

7.1.3 CAM15us

In a fourth series of psychophysical experiments, the brightness perception of 40 different sized, unrelated self-luminous stimuli presented on a display was investigated in a magnitude estimation experiment with 20 observers [90]. The CAM15u model - although constructed using data obtained in different experimental conditions - was able to predict the brightness of the 10° stimuli very well. Furthermore, a significant, hue independent, effect of stimulus size on brightness was found. The effect could be effectively modeled by a simple power function. The impact of the stimulus size on brightness was incorporated into the brightness prediction of the CAM15u model. The predictive performance of the modified brightness prediction, Q_{CAM15us}, was validated using the results obtained in an additional experiment in which twenty observers evaluated the brightness of 62 unrelated self-luminous stimuli having variable size, chromaticities and luminance. Finally, the CAM15u brightness prediction, corrected using the simple power function, Q_{CAM15us}, was compared to that obtained using the size effect transform proposed by Xiao [94] and was found to perform substantially better.

7.2 Some critical reflections

The end-goal of *colour appearance modelling* is to predict the colour appearance of complex 3-dimensional stimuli presented in complex viewing conditions in terms of visual attributes, such as brightness/lightness, colourfulness/saturation/chroma and hue, using the physical (optical) properties of the stimuli and their surroundings as input of the model. Although a complete colour appearance model applicable for complex stimuli in complex viewing conditions is probably years, if not decades, away, the work presented in this dissertation can be considered as one step forward in the prediction of the colour appearance of unrelated self-luminous stimuli.

However, even with the rather severe constraints on the validity of the model which have been mentioned before, the approach adopted in this work, although typical for most colour appearance studies, could have had an impact on the results and the validity of the model.

a) The adaptation state of the observer is known to have an important effect on the colour appearance of a stimulus [102, 103]. Although adaptation under typical viewing conditions and light levels is reported to be generally 90% complete after one minute [102], a balance had to be struck with practical matters, such as the length of a single experimental session (affecting overall experiments duration and observer fatigue). During the experiments, each test stimulus was presented for a fixed period of 15 seconds before presenting the reference stimulus. By using a fixed presentation time the variability in observer response due to differences in adaptation states, and hence adaption time, were minimized. Obviously, some variability still remains due to differences in the time-course of adaptation of each of the observers. In addition, the use of a rather short (stimulus) adaptation time can be argued to be more representative of real-life viewing conditions as people rarely tend to keep their gaze on the object of interest for long.

b) In this dissertation, test stimuli were generated using RGBW led modules. One could question whether the results would hold for stimuli produced using more broadband or even more narrowband (e.g. lasers) spectra. Indeed, some studies report failures in basic colorimetry when applied to more narrowband sources, such as LEDs [104, 105]. However, this is not necessarily a failure of the model itself, but rather a failure of the cone fundamentals used to calculate the cone responses. Currently, the CAM15u model starts from the cone responses determined using the latest cone fundamentals proposed by the CIE (2006), which is also recommended by Csuti and Schanda [104] for narrow band RGB-LEDs.

c) With regards to these cone fundamentals, another issue can be discussed, i.e. the matter of the average age of the observers. In the present study, the age range of the test subjects was deliberately restricted to be between 20 and 32 years, because as people age, the cone fundamentals will change due to yellowing of the lens [45]. Using a restricted age range, limits the variability in the data but also ensures that the chosen cone fundamentals are more representative for the group of observers participating in the experiments. It would be interesting to use different, age dependent cone fundamentals (as proposed by Wold and Farup [106] and the CIE [77]), and to check the performance of the model for observers belonging to several age categories. Finally, as is typical for all colour appearance models, the cone fundamentals are assumed to be those for observers not suffering from any serious colour deficiency, as commonly measured by the Ishihara plate tests or the Farnsworth Munsell 100 Hue test. By using only 2 of the 3 cone responses in the model one could investigate the effect of some types of colour deficiency.

d) In addition to the stimulus presentation and viewing conditions, the experimental method used to derive colour appearance data could have had an impact on the results as well. The magnitude estimation method used in the dissertation requires observers to verbally quantify their perception, which is something quite difficult and which obviously introduces an extra source of error. However, for brightness, the results of the magnitude estimation method were verified using a matching method, in which observers were asked to adjust the luminance of a test stimulus such that it appeared equally bright as a target stimulus. The results obtained with both methods were very consistent. Results for amount of white and hue could also be verified using e.g. a paired comparison, colour discrimination or matching method. However, such extra verification experiments were not performed during the course of this doctoral project.

e) Although CAMs in general, do not aim for the units of their perceptual scales to correspond with a just-noticeable-difference (JND), this would however be a very convenient and informative property. Currently it is unknown whether two stimuli with brightness values of, for example, 31 and 40 are noticeably different. A JND scale would make interpretation of such scales easy, as a difference of 1 would mean the two stimuli are noticeably different for the 'average observer'.

f) The CAM15u model is developed and validated based on data obtained in an experiment using 20 test subjects. The fact that the same observers were used for both development and validation, could have artificially inflated the performance of the CAM15u model in the validation experiment, compared to those calculated for other models proposed in literature. A good way to avoid this issue would have been to ensure that observers participating in the validation experiment had not participated earlier. To check the possible influence of observer overlap between test and validation experiments, the CAM15u model parameters were re-derived 100 times by randomly drawing 10 test subjects from the full panel of 20 observers. For each random draw, the remaining 10 test subjects were then used to validate the CAM15u model with the newly derived model parameters. This way, there is no overlap between the model development and model validation observer panels. This bootstrap method allows to calculate the average (over the 100 draws) model performance, expressed as the coefficient of determination R^2, and its confidence interval for both the CAM15u model applied to model development data and model validation data. The results for brightness (Q), hue quadrature (H) and amount of white (W) are given in Table 7.1.

Table 7.1. Results for brightness (*Q*), hue quadrature (*H*) and amount of white (*W*) of the test of the model performance for CAM15u model applied to the development (Dev.) data and the validation (Val.) data using a bootstrap method.

		Q		H		W	
		Dev.	Val.	Dev.	Val.	Dev.	Val.
$\overline{R^2}$		0.8798	0.8802	0.9902	0.9894	0.8614	0.8582
95%-CI	Lower limit	0.8775	0.8774	0.9898	0.9889	0.8580	0.8542
	Upper limit	0.8821	0.8830	0.9906	0.9899	0.8648	0.8622

As is clear from Table 7.1, the confidence intervals (CI) of the average model performance for the development and validation data overlap. This suggests that the CAM15u model is equally applicable for observers that did not participate in the experiments on which the development of the model was based. However, there could still be observer bias due to the observers being selected from a limited subgroup of the more general population. Note, that in this regard, all experiments performed during this doctoral thesis used a gender balanced group of observers, with one third of them belonging to a group of observers without any connection with the engineering department. However, to further improve the general validity of the model, the test panel could be diversified even more.

g) Although the CAM15u model is shown to work well for the average observer, it cannot account for some peculiar experimental findings. First, about one fifth of the observer's magnitude estimates for hue could not be mapped to the 0-400 hue quadrature scale commonly used in CAMs. However, it should be noted that the commonly used hue quadrature scale has some peculiarities of its own (see Chapter 2), introducing some inconsistencies in the CAM15u model as well (see Section 5.6). Therefore, it might be appropriate to develop an alternative hue attribute. An example might be a more semantically based hue attribute based on the colour categories of Berlin and Kay [107], using red, orange, yellow, green, blue, purple, pink, brown, grey, black and white. A second atypical finding was that about one fifth of the observers reported a substantially lower brightness for very saturated red stimuli. These peculiar results were confirmed in the matching experiment. The exact reason is however unclear, but perhaps these results have been induced by an emotional response to red. Note that there are no experimental arguments indicating that this is an emotional response to red and no such effects were found for other colours.

h) Finally, although the CAM15u model predicts the increase in brightness for more saturated stimuli (cfr. H-K effect) very well and much better than the former existing models, the model does not provide more insights into the where, how and why of the H-K effect.

7.3 Practical applications

As mentioned in the introduction, colour appearance modelling is of interest in numerous disciplines such as colour signalisation (traffic and aviation signal lights, machine control lights,…), printing, computer graphics (colour rendering on displays, colour management software,…), machine vision (image-based automatic inspection and analysis), medical and forensic imaging, photography,… Below, some particularly useful applications of CAM15u are discussed.

7.3.1 Guidelines and standards for signalisation and light pollution

European Standard for variable message traffic signs

In the *European Standard for variable message traffic signs* [108] photometric and colorimetric requirements for these kind of signs are listed. One of the provided requirements deals with the minimum and maximum luminance of the signs, which is dependent not only on the viewing conditions but also on the hue. For example, the maximum luminance for the 'L1' class of signs in a dark surround are 90, 54, 27, 22.5, 21, and 9 cd/m² for white, yellow, green, red, orange, and blue signs, respectively. Also for other classes and surround conditions, the minimum and maximum luminance for white signs is multiplied with 0.6, 0.3, 0.25, 0.233 and 0.1 to obtain the minimum and maximum luminance values for yellow, green, red, orange and blue signs respectively. Note that for orange the values vary, without a given reason, from 0.233 for dark surrounds to 0.387 for bright surrounds.

These hue dependent luminance values and ratios for signs in dark surround conditions can be evaluated using the CAM15u brightness prediction. In Fig.7.1 the 2° x,y chromaticity values of some unrelated self-luminous stimuli used in the visual tests of the doctoral research project, are plotted together with the chromaticity areas for white, yellow, green, red, orange and blue signs providing the best colour distinction according to the standard [108].

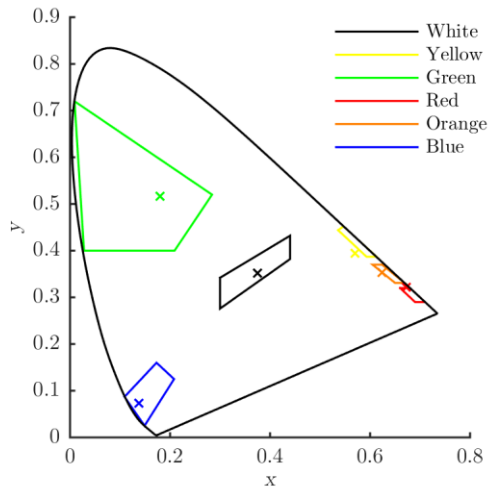


Fig.7.1 x,y chromaticity values of unrelated self-luminous stimuli (crosses) plotted together with the allowed chromaticity areas for white, yellow, green, red, orange and blue signs providing the best colour distinction according to [108].

When one would rely on CAM15u, a standard for LED signs in dark surround conditions should provide luminance values for each characteristic hue such that the CAM15u brightness is similar. In Fig.7.2, the CAM15u brightness predictions of the coloured stimuli with luminance values according to the standard [108] for the 'L1' class of signs in a dark surround are indicated with a circle. The black line represents the CAM15u brightness prediction for the white stimulus having a luminance of 90 cd/m². For green, blue and orange, the standard and the CAM15u prediction clearly correspond well. Red and yellow seem to be slightly underestimated by the standard compared to the model. In Fig. 7.2, the yellow, green, red, orange and blue stimuli having exactly the same brightness as a white stimulus with a luminance of 90 cd/m², are highlighted of with a cross. Their respective luminance values are 28.8, 30.2, 12.7, 19.5 and 10.5 cd/m². These stimuli - with a corresponding ratio of 0.32, 0.34, 0.14, 0.22 and 0.12 - could provide improved maximum luminance values for the 'L1' class of signs in a dark surround.

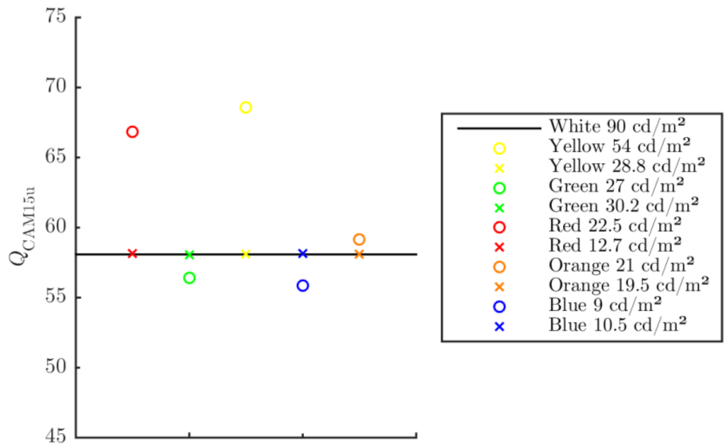


Fig.7.2. CAM15u brightness predictions of unrelated self-luminous stimuli for different 2° luminance values.

Note that, based on the large blue and green area in Fig.7.1, a large variation in saturation is possible for blue and green signs. According to the findings of this doctoral research, large saturation differences cause large brightness differences due to the Helmholtz-Kohlrausch effect. In fact, the standard could be further improved by indicating the maximum luminance values using the chromaticity as an input parameter and using the reverse CAM15u model for brightness.

Dutch guidelines for light pollution and lighting on highways

In the Dutch guideline for light pollution of the NSVV (Nederlandse Stichting Voor Verlichtingskunde), hue independent luminance limits for billboards and for light emission from buildings are provided [109]. For example, a luminance limit of 100 cd/m^2 is provided for billboards between 20 and 50 m^2 in urban areas. The guideline could be improved by making these luminance limits dependent on the hue of the billboard.

The guideline for lighting on highways (ROA – Richtlijnen voor het Ontwerpen van Autosnelwegen), provides rather restricted luminance limits for billboards along highways [109]. In the guideline, different values are given for white, yellow and other coloured billboards. For example, a luminance limit of 500, 250 and 25 cd/m^2 is provided for respectively white, yellow and other colours on billboards smaller than 6 m^2. Again, the guideline could be improved by providing a more extensive hue dependence of the luminance limits and by taking the size effect better into account.

Belgian guidelines for billboards and light pollution

According to the Belgian law (Koninklijk Besluit van 1 december 1975 houdende Algemeen Reglement op de Politie van het Wegverkeer Artikel 80.2) it is forbidden to provide glary or distracting advertising billboards on the public road. In addition, it is forbidden to use red or green billboards (like pharmacy crosses) positioned lower than 7 meter above ground level and less than 75 meter from a traffic light.

The Belgian standard NBN L 18-002 [110] provides maximum luminance values for advertising lighting in showcases and signs. These values are dependent on the 'illuminated' surface and the location (important trade centres or streets). As this standard dates back from the time that billboards were illuminated, it should be revised by providing hue and stimulus size dependent maximum luminance values for self-luminous billboards.

The need for guidelines

With the increased use of LED signs in public areas, the lack of clear, state-of-the-art guidelines for billboards and self-luminous objects is peculiar. An example is given by the Coca Cola billboard on the 'Brouckèreplein' in Brussels (see Fig.7.3). After replacing the neon-lights with a LED display, residents complained about the disturbing brightness. It took more than two years before the display was dimmed. In addition, instead of discussing luminance measurements of the red and white surfaces of the display, the illuminance was measured and used as main parameter. Finally, the display was dimmed to an 'acceptable' level by reducing the illuminance with 20% during the day and with 66% at night. Also in the city of Ghent, questions were raised about the brightness of billboards. Together with the technical service of the city, a master thesis about the brightness of LED billboards has been initiated at the Light&Lighting Laboratory at KU Leuven (not yet finished).

Fig.7.3. The Coca Cola billboard on the Brouckèreplein in Brussels, before dimming (left) and after dimming to an 'acceptable' brightness level during the day (middle) and at night (right). Photograph courtesy of brusselsnieuws.be (left) and V.A.Jacobs (middle and right).

Note that a clear guideline for LED displays could offer energy savings as the coloured displays will often be dimmed to not violate the guidelines' maximum luminance values.

7.3.2 Glare

The assessment of the discomfort glare of a luminaire is in fact related to the perceived brightness of the luminaire. Having a model that is capable of predicting the brightness of a (uniform) light source, one can construct a corresponding "glariness scale" by doing some additional experiments with a test panel. However, LED luminaires often constitute a non-uniform stimulus with high and low luminance regions. For these kind of stimuli, a CAM is needed which can predict the brightness of non-uniform stimuli.

7.4 Future research

There is still a lot of research needed before a general CAM applicable for all possible viewing conditions can be established. First, CAM15u should be extended for luminous backgrounds. Therefore visual experiments with a stimulus seen against an achromatic background characterized by its colour temperature and even seen against a coloured background are needed. In these investigations, adaptation and colour contrast will be key aspects. A PhD project about this topic is just started in Argentina (University of Tucumán). Based on the results of both this thesis and the research in Argentina, luminous backgrounds will also be tackled at the Light&Lighting Laboratory at KU Leuven in order to extend the CAM15u model. The absolute spectral radiance of both stimulus and background would act as the only input parameters.

Second, the research should be extended to related colours by investigating the H-K effect for these colours, incorporating some of the steps of CAM15u in the existing CAMs for related colours (e.g. by starting with LMS signals in CIECAM02 and by evaluation of the amount of white),...

Third, other factors such as adaptation, stimulus shape, stimulus position in the visual field, viewing distance,... could be investigated. For example, a stimulus could be evaluated after a viewing time of 1 second on one hand and after 1 minute on the other hand, one could also use non-uniform stimuli or characters (like pharmacy crosses) or one could select the same FOV of 10° but at different viewing distances,...

Fourth, the brightness scale could be improved by investigating just noticeable differences (JNDs). In this way, the brightness scale could be rescaled such that an interval of one unit brightness would correspond to one JND.

Fifth, a CAM applicable to people of different ages could be developed. Such a CAM would be able to predict approximate colour attributes for observers depending on their age and visual impairments and eye pathology. When studying the effect of age on the colour appearance, the observers should be carefully selected as visual experiments with observers above 60 seem to be difficult. Furthermore, the variability between elderly observers will be quite large compared to younger observers.

Sixth, a better 'colour difference formula' between stimuli could be developed. Indeed, when a good CAM is available, the correlates of the CAM can be used as basic scales on which colour difference is evaluated. By asking observers to judge the colour difference between stimuli that vary in brightness, hue and colourfulness or amount of white, the weight of each attribute in the perception of colour could be found. This improves the development of a meaningful colour difference formula, a very useful application for industrial purposes.

And finally, field tests could be developed to investigate the colour appearance of a complex scene, room, public space,... For these kind of stimuli, 'stimulus' and background are not uniform any more, and even the distinction between stimulus and background becomes irrelevant. Only by performing visual experiments in a real, complex environment, the gap between colour science, lighting design and architecture can hopefully be reduced.

Chapter 8

REFERENCES

1. CIE, *International Lighting Vocabulary*, in *CIE S 017/E:2011*. 2011, CIE Central Bureau: Vienna, Austria.
2. Hunt, R.W.G., *A model of colour vision for predicting colour appearance in various viewing conditions.* Color Research & Application, 1987. **12**(6): p. 297-314.
3. Hunt, R.W.G., *A model of colour vision for predicting colour appearance.* Color Research & Application, 1982. **7**(2): p. 95-112.
4. Hunt, R.W.G. and M.R. Pointer, *A colour-appearance transform for the CIE 1931 standard colorimetric observer.* Color Research & Application, 1985. **10**(3): p. 165-179.
5. Nayatani, Y., et al., *A nonlinear color-appearance model using estévez-hunt-pointer primaries.* Color Research & Application, 1987. **12**(5): p. 231-242.
6. Nayatani, Y., K. Takahama, and H. Sobagaki, *Prediction of color appearance under various adapting conditions.* Color Research & Application, 1986. **11**(1): p. 62-71.
7. Fairchild, M.D., *Chromatic adaptation to image displays.* TAGA, 1992. **2**: p. 803-824.
8. Fairchild, M.D., *Refinement of the RLAB color space.* Color Research & Application, 1998. **21**: p. 338-346.
9. Luo, M.R., et al., *Quantifying colour appearance. Part I. Lutchi colour appearance data.* Color Research & Application, 1991. **16**(3): p. 166-180.
10. Luo, M.R., et al., *Quantifying colour appearance. Part II. Testing colour models performance using lutchi colour appearance data.* Color Research & Application, 1991. **16**(3): p. 181-197.
11. Luo, M.R. and R.W.G. Hunt, *The structure of the CIE 1997 Colour Appearance Model (CIECAM97s).* Color Research & Application, 1998. **23**(3): p. 138-146.
12. CIE, *A colour appearance model for colour management systems: CIECAM02*, in *Publication 159.* 2004, CIE Central Bureau: Vienna, Austria. p. 19.
13. Kim, M.H., T. Weyrich, and J. Kautz. *Modeling human color perception under extended luminance levels.* in *International Conference on Computer Graphics and Interactive Techniques:*

SIGGRAPH '09 2009. New Orleans, Louisiana, USA: ACM New York, NY, USA.

14. Kunkel, T. and E. Reinhard, *A neurophysiology-inspired steady-state color appearance model.* J. Opt. Soc. Am. A, 2009. **26**(4): p. 776-782.
15. Hunt, R.W.G., *Revised colour-appearance model for related and unrelated colours.* Color Research & Application, 1991. **16**(3): p. 146-165.
16. Hunt, R.W.G., *Measuring colour*. Third ed. 1998, Kingston-upon-Thames, England: Fountain Press. 239-246.
17. Fu, C., et al., *An investigation of colour appearance for unrelated colours under photopic and mesopic vision.* Color Research & Application, 2012. **37**(4): p. 238-254.
18. Rames, C. *CIECAM02 Plug-in*. Available from: https://sites.google.com/site/clifframes/ciecam02plugin.
19. Gill, G. *Argyll Color Management System*. Available from: http://www.argyllcms.com/.
20. Moroney, N., et al., *The CIECAM02 Color Appearance Model*, in *10th Color Imaging Conference, IS&T and SID*. 2002: Scottsdale, Arizona.
21. Galbraith, R. *Microsoft and Canon join forces on Vista's Windows Color System*. 2005; Available from: http://www.robgalbraith.com/content_page4aeb.html?cid=7-7886-8011.
22. Luo, M.R. and C. Li, *CIECAM02 and Its Recent Developments*, in *Advanced Color Image Processing and Analysis*. 2013, Springer.
23. Withouck, M., et al., *Brightness perception of unrelated self-luminous colors.* Journal of the Optical Society of America A, 2013. **30**(6): p. 1248-1255.
24. Fairchild, M.D., *Color Appearance Models*. Third ed. 2013, Chichester, UK: Wiley-IS&T.
25. Kotani, T., et al. *Studies on the lighting method to achieve both energy saving and the comfort using luminance based lighting design*. in *CIE 2010 Lighting Quality and Energy Efficiency*. 2010. Vienna, Austria.
26. Nakamura, Y. *Quantitative expression of contrast in a measured luminance image*. in *10th Lux Europa*. 2005. Berlin, Germany.
27. Hunt, R.W.G., *Hue shifts in unrelated and related colours.* Color Research & Application, 1989. **14**(5): p. 235-239.
28. Fu, C., et al. *Colour-appearance modelling for unrelated colour under photopic and mesopic vision*. in *Proc. IS&T/SID 15th Color Imaging Conference*. 2007. New Mexico.
29. Hurvich, L.M. and D. Jameson, *A Psychophysical Study of White. III. Adaptation as Variant.* Journal of the Optical Society of America, 1951. **41**(11): p. 787-800.

30. Valeton, J.M. and D. van Norren, *Light adaptation of primate cones: An analysis based on extracellular data.* Vision Research, 1983. **23**(12): p. 1539-1547.
31. Hunt, R.W.G. and M.R. Pointer, *Measuring colour*. Fourth ed. Wiley-IS&T Series in Imaging Science and Technology, ed. M.A. Kriss. 2011, West Sussex, United Kingdom: John Wiley & Sons Ltd.
32. Stiles, W.S. and J.M. Burch, *N.P.L. Colour-matching Investigation: Final Report (1958).* Optica Acta: International Journal of Optics, 1959. **6**(1): p. 1-26.
33. Estévez, O., *On the fundamental data-base of normal and dichromatic colour vision*. 1979, Leiden.
34. Padgham, C.A. and S.C.H. Rowe. *A Mass Colour Scaling Experiment*. in *AIC Colour 73*. 1973. Bristol, England: Adam Hilger.
35. Adams, E.Q., *X-Z Planes in the 1931 I.C.I. System of Colorimetry1.* Journal of the Optical Society of America, 1942. **32**(3): p. 168-173.
36. Bartleson, C.J., *Measures of Brightness and Lightness.* Die Farbe, 1980. **28**: p. 132-148.
37. Fu, C., et al., *Investigation of unique hue setting changes with ageing.* Chinese Optics Letters, 2011. **9**(5): p. 053301.
38. Hunt, R.W.G., C.J. Li, and M.R. Luo, *Dynamic cone response functions for models of colour appearance.* Color Research & Application, 2003. **28**(2): p. 82-88.
39. Cornsweet, T.N., *Visual perception*. 1970, New York: Academic.
40. Scheffé, H., *An Analysis of Variance for Paired Comparisons.* Journal of the American Statistical Association, 1952. **47**(259): p. 381-400
41. Leloup, F., et al. *Integration of Multiple Cues for Visual Gloss Evaluation*. in *Predicting Perceptions: Proceedings of the 3rd International Conference on Appearance*. 2012. Edinburgh, UK.
42. Leloup, F.B., et al., *Geometry of illumination, luminance contrast, and gloss perception.* J. Opt. Soc. Am. A, 2010. **27**(9): p. 2046-2054.
43. ASTM International, *Standard Test Method for Unipolar Magnitude Estimation of Sensory Attributes*. 2012: West Conshohocken, United States.
44. Gescheider, G.A., *Psychophysical scaling.* Annu Rev Psychol, 1988. **39**: p. 169-200.
45. CIE, *Guidance towards Best Practice in Psychophysical Procedures Used when Measuring Relative Spatial Brightness*. 2014, CIE: Vienna, Austria.
46. Weller, S.C. and A.K. Romney, *Systematic Data Collection* 1988: Sage Publications, Inc.
47. Farnsworth, D., *The Farnsworth-Munsell 100-hue test for examination of colour discrimination*. 1957, Munsell Color Co.: Baltimore.
48. Leloup, F.B., et al., *Luminance-based specular gloss characterization.* J. Opt. Soc. Am. A, 2011. **28**(6): p. 1322-1330.

49. Fotios, S.A. and C. Cheal, *A comparison of simultaneous and sequential brightness judgements.* Lighting Research & Technology, 2010. **42**(2): p. 183-197.
50. García, P.A., et al., *Measurement of the relationship between perceived and computed color differences.* J. Opt. Soc. Am. A, 2007. **24**(7): p. 1823-1829.
51. Sanders, C.L. and G. Wyszecki. *Correlate for brightness in terms of CIE color matching data*. in *CIE Proceedings 15th Session*. 1963. Vienna: CIE Central Bureau, Paris.
52. Nayatani, Y., *Simple estimation methods for the Helmholtz—Kohlrausch effect.* Color Research & Application, 1997. **22**(6): p. 385-401.
53. CIE, *Supplementary System of Photometry*. 2011, CIE Central Bureau: Vienna, Austria.
54. CIE, *Testing of Supplementary Systems of Photometry*. 2001, CIE Central Bureau: Vienna, Austria.
55. Wyszecki, G. and W.S. Stiles, *Color Science*. 1982, Wiley: Canada. p. 410.
56. Guth, S.L., N.J. Donley, and R.T. Marrocco, *On luminance additivity and related topics.* Vision Research, 1969. **9**(5): p. 537-IN1.
57. Guth, S.L. and H.R. Lodge, *Heterochromatic additivity, foveal spectral sensitivity, and a new color model.* J. Opt. Soc. Am., 1973. **63**(4): p. 450-462.
58. Boynton, R.M. and P.K. Kaiser, *Vision: the additivity law made to work for heterochromatic photometry with bipartite fields.* Science, 1968. **161**(3839): p. 366-8.
59. Wagner, G. and R.M. Boynton, *Comparison of Four Methods of Heterochromatic Photometry.* J. Opt. Soc. Am., 1972. **62**(12): p. 1508-1515.
60. Kaiser, P.K. and G. Wyszecki, *Additivity failures in heterochromatic brightness matching.* Color Research & Application, 1978. **3**(4): p. 177-182.
61. Stevens, S.S., *On the psychophysical law.* Psychological Review, 1957. **64**(3): p. 153-81.
62. Yaguchi, H. and M. Ikeda, *Subadditivity and superadditivity in heterochromatic brightness matching.* Vision Res, 1983. **23**(12): p. 1711-8.
63. Guth, S.L. *ATD01 model for color appearances, color differences and chromatic adaptation*. in *9th Congress of the International Colour Association, Proceedings of SPIE Vol. 4421*. 2002.
64. Capilla, P., et al., *On saturation and related parameters following Guth's ATD colour-vision model.* Color Research & Application, 2001. **26**(4): p. 305-321.

65. Fairchild, M.D., *Color Appearance Models*. Second ed. Wiley-IS&T Series in Imaging Science and Technology, ed. M.A. Kriss. 2005, West Sussex, England: John Wiley & Sons Ltd.
66. Sagawa, K. and K. Takeichi, *System of mesopic photometry for evaluating lights in terms of comparative brightness relationships.* J. Opt. Soc. Am. A, 1992. **9**(8): p. 1240-1246.
67. Nakano, Y., *A model for brightness perception and its application to individual data.* Jpn. J. Opt., 1992. **21**: p. 705-716.
68. Palmer, D.A., *Standard Observer for Large-Field Photometry at any Level.* J. Opt. Soc. Am., 1968. **58**(9): p. 1296-1298.
69. Koo, B. and Y. Kwak, *Color appearance and color connotation models for unrelated colors.* Color Research & Application, 2015. **40**(1): p. 40-49.
70. CIE, *Colorimetry*. 2004, CIE Central Bureau: Vienna, Austria.
71. Field, A., *Discovering statistics using SPSS*. Third ed. 2009: SAGE.
72. Withouck, M., et al., *Predicting the brightness of unrelated self-luminous stimuli.* Optics Express, 2014. **22**(13): p. 16298-16309.
73. *IBM SPSS Statistics for Windows*. 2012, IBM Corp. : Armonk, NY.
74. Fotios, S., K. Houser, and C. Cheal, *Counterbalancing needed to avoid bias in side-by-side brightness matching tasks.* Leukos, 2008. **4**(4).
75. Tversky, A. and D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases.* Science, 1974. **185**(4157): p. 1124-31.
76. Withouck, M., et al., *Experimental driven modelling of the color appearance of unrelated self-luminous stimuli: CAM15u.* Optics Express, 2015. **23**(9): p. 12045-12064.
77. CIE, *Fundamental chromaticity diagram with physiological axes - part 1.* 2006, CIE: Vienna, Austria.
78. Stockman, A. and L.T. Sharpe, *The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype.* Vision Res, 2000. **40**(13): p. 1711-37.
79. Stockman, A., L.T. Sharpe, and C. Fach, *The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches.* Vision Res, 1999. **39**(17): p. 2901-27.
80. Smet, K.A.G., G. Deconinck, and P. Hanselaer, *Chromaticity of unique white in object mode.* Optics Express, 2014. **22**(21): p. 25830-25841.
81. Vos, J.J. and P.L. Walraven, *On the derivation of the foveal receptor primaries.* Vision Research, 1971. **11**(8): p. 799-818.
82. Carroll, J., J. Neitz, and M. Neitz, *Estimates of L:M cone ratio from ERG flicker photometry and genetics.* J Vis, 2002. **2**(8): p. 531-42.
83. Wyszecki, G. and W.S. Stiles, *Color Science*. Second ed. 1982, Canada: John Wiley & Sons Inc.

84. Schuchard, R.A., *Review of colorimetric methods for developing and evaluating uniform CRT display scales.* Optical Engineering, 1990. **29**(4): p. 378-384.
85. Brill, M.H. and R.C. Carter, *Does lightness obey a log or a power law? Or is that the right question?* Color Research & Application, 2014. **39**(1): p. 99-101.
86. !!! INVALID CITATION !!!
87. Withouck, M., et al. *Size and color do matter when predicting brightness*. in *Vision Sciences Society 14th Annual Meeting*. 2014. St. Pete Beach, Florida.
88. CIE, *Effect of stimulus size on colour appearance*. 2011, CIE Central Bureau: Vienna.
89. Xiao, K., et al., *Colour appearance of room colours.* Color Research & Application, 2010. **35**(4): p. 284-293.
90. Withouck, M., K.A.G. Smet, and P. Hanselaer, *Brightness prediction of different sized unrelated self-luminous stimuli.* Optics Express, 2015. **23**(10): p. 13455-13466.
91. Ronchi, L.R., *On the dependence of brightness on target size*, in *Fondazione Giorgio Ronchi*. 2002. p. 79-95.
92. Gombos, K. and J. Schanda. *Interrelationship between size and brightness dimensions of appearance*. in *CIE Expert Symposium on Visual Appearance*. 2006. Paris: University of Pannonia.
93. Xiao, K., *Colour Appearance Assessment for Dissimilar Sizes*. 2006, University of Derby.
94. Xiao, K., M.R. Luo, and C. Li, *Colour size effect modelling.* Color Research & Application, 2012. **37**(1): p. 4-12.
95. Xiao, K., et al., *Investigation of colour size effect for colour appearance assessment.* Color Research & Application, 2011. **36**(3): p. 201-209.
96. Luo, M., G. Cui, and M. Georgoula, *Colour difference evaluation for white light sources.* Lighting Research and Technology, 2014.
97. Fotios, S.A. and C. Cheal, *Lighting for subsidiary streets: investigation of lamps of different SPD. Part 2—Brightness.* Lighting Research and Technology, 2007. **39**(3): p. 233-249.
98. Loe, L., K.P. Mansfield, and E. Rowlands, *Appearance of lit environment and its relevance in lighting design: Experimental study.* Lighting Research and Technology, 1994. **26**(3): p. 119-133.
99. Dubois, M.C. *Effect of glazing types on daylight quality in interiors: conclusions from three scale model studies*. in *Experiencing Light*. 2009. Eindhoven.
100. Cowdroy, R.M., *A study of the scalar effect in the use of models for glare assessment*. 1972, University of Sydney: Sydney.
101. Lau, J.J.H., *Use of scale models for appraising lighting quality.* Lighting Research and Technology, 1972. **4**(4): p. 254-262.

102. Fairchild, M.D. and L. Reniff, *Time course of chromatic adaptation for color-appearance judgments.* J. Opt. Soc. Am. A, 1995. **12**(5): p. 824-833.
103. Shevell, S., *The Time Course of Chromatic Adaptation.* Color Research & Application, 2001. **26**: p. 170-173.
104. Csuti, P. and J. Schanda, *Colour matching experiments with RGB-LEDs.* Color Research & Application, 2008. **33**(2): p. 108-112.
105. Csuti, P. and J. Schanda, *Colour Matching Based on Fundamental Spectral Sensitivity Functions*, in *ISCC/CIE Expert Symposium*. 2006: Ottawa, Canada.
106. Wold, J.H. and I. Farup, *Age- and field-size-parameterized calculations of physiologically significant XYZ colour-matching functions*, in *Color, Light & Appearance Week at NIST*. 2014: Gaithersburg, Maryland, USA.
107. Berlin, B. and P. Kay, *Basic Color Terms: Their Universality and Evolution*. 1969: University of California Press.
108. CEN, *Road vertical signs - Variable message traffic signs*, in *EN 12966:2014*. 2014, BSI.
109. NSVV, *Richtlijn lichthinder*. 2014, GVO drukkers & vormgevers: Ede, Nederland.
110. NBN, *Aanbevelingen voor bijzondere gevallen van openbare verlichting*. 1988.

CURRICULUM VITAE

General:

Withouck Martijn, 23/03/1987 (Kortrijk)

Education:

Bachelor degree in Specific Teacher Training (SLO), CVO VIVO Kortrijk, 2014.

Master degree in Physics and astronomy (minor economics and business), Ghent University, 2010.

Bachelor degree in Physics and astronomy, Ghent University, 2008.

Mathematics-Science, Sint-Aloysiuscollege, Menen, 2005.

Experience:

01/06/2015 – present: Research Engineer at FLIR Intelligent Transportation Systems.

17/10/2011 – 31/05/2015: PhD Researcher at the Light&Lighting Laboratory, KU Leuven.

01/09/2010 – 16/10/2011: Research fellow at the Light&Lighting Laboratory, Catholic University College Ghent.

LIST OF PUBLICATIONS

Articles in internationally reviewed academic journals

Jacobs, V., Lenseigne, B., **Withouck, M.**, Hanselaer, P., Jonker, P. (2013). Color sensitivity of the multi-exposure HDR imaging process. *Advanced Optical Technologies, 2* (2), 159-169.

Withouck, M., Smet, K., Ryckaert, W., Pointer, M., Deconinck, G., Koenderink, J., Hanselaer, P. (2013). Brightness perception of unrelated self-luminous colors. *Journal of the Optical Society of America A, Optics, Image Science and Vision, 30* (6), 1248-1255.
-> selected by the Editors for publication in the Virtual Journal for Biomedical Optics (VJBO)

Withouck, M., Smet, K., Ryckaert, W., Deconinck, G., Hanselaer, P. (2014). Predicting the brightness of unrelated self-luminous stimuli. *Optics Express, 22* (13), 16298-16309.
-> selected by the Editors for publication in the Virtual Journal for Biomedical Optics (VJBO)

Withouck, M., Smet, K., Ryckaert, W., Hanselaer, P. (2015). Experimental driven modelling of the color appearance of unrelated self-luminous stimuli: CAM15u. *Optics Express, 23* (9), 12045-12064.

Withouck, M., Smet, K., Hanselaer, P. (2015). Brightness prediction of different sized unrelated self-luminous stimuli. *Optics Express, 23* (10), 13455-13466.

Papers at international scientific conferences and symposia, published in full in proceedings

Withouck, M., Ryckaert, W., Smet, K., Deconinck, G., Hanselaer, P. (2012). Colour Appearance Modelling for Self-luminous Colours. *Proceedings of the 3rd International Conference on Appearance. International Conference on Appearance: Predicting Perceptions*. Edinburgh, Scotland, 17-19 April 2012 (pp. 147-149). Edinburgh, UK – Oral presentation

Withouck, M., Ryckaert, W., De Wilde, P., Deconinck, G., Hanselaer, P. (2012). Age Related Colour Appearance Differences for Unrelated Self-Luminous Colours. *PROCEEDINGS of CIE 2012 „Lighting Quality and Energy Efficiency".* Lighting Quality and Energy Efficiency. Hangzhou, CHINA, 19-21 September 2012 (art.nr. OP50) (pp. 385-390) – Oral presentation

Withouck M., Wattez J., Ryckaert W.R., Smet K., Deconinck G., Hanselaer P. (2013) Colour appearance differences for self-luminous unrelated colours with 2º and 10º field of view. "*Proceedings of 12th International AIC Colour Congress: Bringing Colour to Life".* Gateshead, UK, 8-12 July 2013 (pp.1289-1292) – Poster presentation

Leyre, S., **Withouck, M.**, Durinck, G., Hofkens, J., Deconinck, G., Hanselaer, P. (2014). Modelling Fluorescent Materials with a Spectral Overlap between Excitation and Emission Spectrum. *Proceedings 2nd International Conference on Photonics, Optics and Laser Technology*. International Conference on Photonics, Optics and Laser Technology. Lisbon, 7-9 January 2014 (pp. 33-38).

Withouck, M., Smet, K., Ryckaert, W., Hanselaer, P. (2015). A new colour appearance model for unrelated colours. *CIE 2015.* Manchester, UK, 29 June – 3 July 2015. – Oral presentation

Meeting abstracts, presented at international scientific conferences and symposia, published or not published in proceedings or journals

Withouck, M., Smet, K., Ryckaert, W., Wattez, J., Deconinck, G., Hanselaer, P. (2014). Size and color do matter when predicting brightness. *Vision Science Society: Abstracts 2014.* Vision Sciences Society 14th Annual Meeting. St. Pete Beach, Florida, 16-21 May 2014, 32. – Poster presentation